GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 23, 2004, 13:57:29 ; Search time 179 Seconds
(without alignments)
474.659 Million cell updates/sec

Title: US-10-798-923A-36
Perfect score: 20
Sequence: 1 agtaacatctatgtttggtt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 3399856

Minimum DB seq length: 0
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	14.4	72.0	25	8	ACI60534	Aci60534 Human mic
	2	13.8	69.0	20	3	AAA61982	Aaa61982 Human MEK
	3	13.8	69.0	33	3	AAA40172	Aaa40172 H. pylori
	4	13.8	69.0	33	4	AAF88123	Aaf88123 H. pylori
	5	13.8	69.0	33	4	AAF88066	Aaf88066 H. pylori
	6	13.8	69.0	43	3	AAA13992	Aaa13992 Geranylge
c	7	13.6	68.0	24	6	ABL99963	Abl99963 HOMO 2-hy
c	8	13.6	68.0	28	4	AAF77839	Aaf77839 Glycerol
c	9	13.4	67.0	25	8	ACI17115	Aci17115 Human mic
c	10	13.4	67.0	31	2	AAX06450	Aax06450 Human bia
	11	13.2	66.0	24	2	AAT89524	Aat89524 Bloom's s
	12	13.2	66.0	65	6	ABZ29491	Abz29491 Candida g
	13	12.8	64.0	17	7	ACD50544	Acd50544 HBV hamme
	14	12.8	64.0	17	7	ACD51915	Acd51915 HBV inozy
c	15	12.8	64.0	25	2	AAQ21867	Aaq21867 Probe 179
	16	12.8	64.0	25	8	ACI60535	Aci60535 Human mic
	17	12.8	64.0	25	8	ACI01532	Aci01532 Human mic
	18	12.8	64.0	25	8	ACH61845	Ach61845 DNA targe
	19	12.8	64.0	25	8	ACH57384	Ach57384 DNA targe
	20	12.8	64.0	30	4	AAS09822	Aas09822 Oat Beta-
	21	12.8	64.0	34	4	AAF83283	Aaf83283 Human Chk
c	22	12.8	64.0	43	6	ABZ27765	Abz27765 Candida e
c	23	12.8	64.0	47	3	AAZ67540	Aaz67540 Human map
c	24	12.8	64.0	51	5	ABL00484	Abl00484 Human sil
c	25	12.8	64.0	58	3	AAZ44622	Aaz44622 Newcastle
c	26	12.8	64.0	74	4	AAF83273	Aaf83273 S. cerevi
	27	12.6	63.0	20	7	ACC73349	Acc73349 M marinum
c	28	12.6	63.0	24	2	AAX02618	Aax02618 S. aureus
c	29	12.6	63.0	25	8	ACK28029	Ack28029 Human mic
c	30	12.6	63.0	25	8	ACI29919	Aci29919 Human mic
c	31	12.6	63.0	25	8	ACI36381	Aci36381 Human mic
	32	12.6	63.0	25	8	ACI65204	Aci65204 Human mic
	33	12.6	63.0	25	8	ACK00121	Ack00121 Human mic
	34	12.6	63.0	25	8	ACI27083	Aci27083 Human mic
c	35	12.6	63.0	25	8	ACI35745	Aci35745 Human mic
c	36	12.6	63.0	25	8	ACI25132	Aci25132 Human mic
	37	12.6	63.0	25	8	ACI35606	Aci35606 Human mic
c	38	12.6	63.0	27	2	AAZ18573	Aaz18573 Primer fo
c	39	12.6	63.0	27	3	AAA80480	Aaa80480 ASTH1 pol
c	40	12.6	63.0	27	7	ACC72239	Acc72239 Forward A
	41	12.6	63.0	34	2	AAQ37473	Aaq37473 Sequence
c	42	12.6	63.0	50	6	ABK91114	Abk91114 50 bp spa
	43	12.6	63.0	50	6	ABZ00880	Abz00880 Human leu
c	44	12.6	63.0	50	6	ABZ03294	Abz03294 Human leu
	45	12.6	63.0	55	6	ABZ28986	Abz28986 Candida g
	46	12.6	63.0	60	3	AAA37505	Aaa37505 Yeast acy
c	47	12.6	63.0	60	6	ABN46014	Abn46014 Human spl
c	48	12.6	63.0	63	3	AAC30234	Aac30234 Human sec
	49	12.6	63.0	65	6	ABN30788	Abn30788 Rat splic
	50	12.6	63.0	65	6	ABN50991	Abn50991 Mouse spl
	51	12.6	63.0	74	7	ACD95453	Acd95453 Human col
	52	12.6	63.0	74	7	ACD93726	Acd93726 Human col
	53	12.6	63.0	80	4	AAS23680	Aas23680 Tetracycl
c	54	12.4	62.0	20	6	ABQ93193	Abq93193 T. tausch
	55	12.4	62.0	25	8	ACI98895	Aci98895 Human mic
	56	12.4	62.0	25	8	ACI61996	Aci61996 Human mic
	57	12.4	62.0	25	8	ACI99520	Aci99520 Human mic
	58	12.4	62.0	25	8	ACH57128	Ach57128 DNA targe
	59	12.4	62.0	25	8	ACH52746	Ach52746 DNA targe
	60	12.4	62.0	25	8	ACH61209	Ach61209 DNA targe
	61	12.4	62.0	47	3	AAZ68365	Aaz68365 Human map
	62	12.4	62.0	50	6	ABZ05083	Abz05083 Human leu
c	63	12.4	62.0	51	5	ABL00324	Abl00324 Human sil
	64	12.4	62.0	60	6	ABN40147	Abn40147 Human spl
c	65	12.4	62.0	60	6	ABN40945	Abn40945 Human spl
	66	12.4	62.0	65	6	ABN31558	Abn31558 Rat splic
	67	12.2	61.0	20	2	AAT05166	Aat05166 Human cyt
	68	12.2	61.0	21	4	AAD21805	Aad21805 Human ATM
	69	12.2	61.0	23	6	ABV73143	Abv73143 CYP2D6 ge
	70	12.2	61.0	23	6	ABV74937	Abv74937 CYP2D6 ge
	71	12.2	61.0	23	6	ABA00287	Aba00287 CYP2D6 C1
	72	12.2	61.0	23	7	AAD54276	Aad54276 CYP2D6 mu
	73	12.2	61.0	23	7	ABZ75812	Abz75812 CYP2D6 ge
	74	12.2	61.0	23	7	AAD50999	Aad50999 CYP2D6 ge
	75	12.2	61.0	23	7	ABZ20545	Abz20545 Human CYP
	76	12.2	61.0	23	7	ABZ23163	Abz23163 PCR prime
	77	12.2	61.0	23	7	ABV72596	Abv72596 PCR prime
	78	12.2	61.0	23	7	AAL53892	Aal53892 Gastroeso
	79	12.2	61.0	23	7	AAD51935	Aad51935 CYP2D6 mu
	80	12.2	61.0	23	7	ABZ23451	Abz23451 Primer us
	81	12.2	61.0	23	7	ABZ24378	Abz24378 Human cyt
	82	12.2	61.0	23	7	AAD47713	Aad47713 CYP2D6 C1
	83	12.2	61.0	23	7	ABV76252	Abv76252 Cytochrom
	84	12.2	61.0	23	8	ACF05589	Acf05589 Cytochrom
	85	12.2	61.0	23	8	ACC79740	Acc79740 CYP2D6 C1
	86	12.2	61.0	23	8	ACC83576	Acc83576 Cytochrom
	87	12.2	61.0	23	9	ACF04628	Acf04628 Human CYP
	88	12.2	61.0	24	3	AAA27864	Aaa27864 Serum and
c	89	12.2	61.0	25	2	AAT36471	Aat36471 CFTR gene
	90	12.2	61.0	25	8	ACK08374	Ack08374 Human mic
	91	12.2	61.0	25	8	ACK13225	Ack13225 Human mic
	92	12.2	61.0	25	8	ACI19523	Aci19523 Human mic
	93	12.2	61.0	26	2	AAX59988	Aax59988 Oligonucl
	94	12.2	61.0	32	7	ABQ83378	Abq83378 Human NR1
c	95	12.2	61.0	33	2	AAQ87760	Aaq87760 Human aux
c	96	12.2	61.0	33	2	AAT48338	Aat48338 Primer fo

c	97	12.2	61.0	33	2	AAT26950
c	98	12.2	61.0	34	2	AAT00595
	99	12.2	61.0	38	4	AAH48963
	100	12.2	61.0	41	6	ABZ44088
	101	12.2	61.0	41	6	ABZ50030
c	102	12.2	61.0	41	6	ABL61320
	103	12.2	61.0	42	4	AAD17242
c	104	12.2	61.0	44	4	AAD17241
	105	12.2	61.0	45	4	AAF62754
	106	12.2	61.0	45	6	ABK91388
	107	12.2	61.0	50	6	ABZ04349
	108	12.2	61.0	50	6	ABZ06504
c	109	12.2	61.0	50	6	ABZ06894
	110	12.2	61.0	50	6	ABZ02895
c	111	12.2	61.0	51	4	AAL28736
	112	12.2	61.0	51	4	AAI73673
	113	12.2	61.0	51	4	AAH38776
	114	12.2	61.0	56	2	AAT21822
c	115	12.2	61.0	60	6	ABN36821
c	116	12.2	61.0	60	6	ABN50252
c	117	12.2	61.0	65	6	ABN31593
c	118	12.2	61.0	72	6	ABA04055
	119	12.2	61.0	75	3	AAA56067
c	120	12.2	61.0	78	3	AAC10957
c	121	12	60.0	24	5	AAH55982
	122	12	60.0	25	3	AAC71018
	123	12	60.0	25	3	AAC70961
	124	12	60.0	25	3	AAC70964
	125	12	60.0	25	3	AAC71012
	126	12	60.0	25	3	AAC70991
	127	12	60.0	25	8	ACI83535
	128	12	60.0	25	8	ACI40063
c	129	12	60.0	25	8	ACI77257
	130	12	60.0	25	8	ACI31661
	131	12	60.0	25	8	ACI83534
	132	12	60.0	25	8	ACI55181
c	133	12	60.0	25	8	ACI74139
	134	12	60.0	25	8	ACI93584
c	135	12	60.0	25	8	ACI16296
	136	12	60.0	28	5	AAF23976
	137	12	60.0	32	4	AAD12815
	138	12	60.0	42	3	AAZ61102
	139	12	60.0	42	3	AAZ61091
	140	12	60.0	45	2	AAT07616
	141	12	60.0	45	2	AAT02624
c	142	12	60.0	50	2	AAQ12736
c	143	12	60.0	51	3	AAA77164
c	144	12	60.0	51	3	AAA77165
c	145	12	60.0	51	4	AAL27334
	146	12	60.0	51	6	ABZ47058
	147	12	60.0	52	4	ABL56759
	148	12	60.0	57	2	AAV15677
c	149	12	60.0	58	2	AAV79293
	150	12	60.0	60	6	ABN38091

Aat26950	Human cyt
Aat00595	GM-CSF am
Aah48963	Human CFT
Abz44088	Human NDU
Abz50030	Human NDU
Abl61320	Human B l
Aad17242	Human CFT
Aad17241	Human CFT
Aaf62754	Primer us
Abk91388	PCR prime
Abz04349	Human leu
Abz06504	Human leu
Abz06894	Human leu
Abz02895	Human leu
Aal28736	Human SNP
Aai73673	Human sil
Aah38776	Human SNP
Aat21822	Human gen
Abn36821	Human spl
Abn50252	Human spl
Abn31593	Rat splic
Aba04055	HIV gag z
Aaa56067	Inhibitor
Aac10957	Human sec
Aah55982	Human SCN
Aac71018	Single nu
Aac70961	Single nu
Aac70964	Single nu
Aac71012	Single nu
Aac70991	Single nu
Aci83535	Human mic
Aci40063	Human mic
Aci77257	Human mic
Aci31661	Human mic
Aci83534	Human mic
Aci55181	Human mic
Aci74139	Human mic
Aci93584	Human mic
Aci16296	Human mic
Aaf23976	Human 5-H
Aad12815	Solanum t
Aaz61102	Reverse P
Aaz61091	Reverse P
Aat07616	RT-PCR pr
Aat02624	Primer 26
Aaq12736	Mip gene
Aaa77164	Human clo
Aaa77165	Human clo
Aal27334	Human SNP
Abz47058	Human ATP
Abl56759	Nucleotid
Aav15677	PCR prime
Aav79293	Staphyloc
Abn38091	Human spl

ALIGNMENTS

RESULT 1
ACI60534
ID ACI60534 standard; DNA; 25 BP.
XX
AC ACI60534;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human microarray DNA oligonucleotide SEQ ID NO 60525.
XX
KW EST; ss; probe; expressed sequence tag; microarray; gene expression;
KW genetic variation; biallelic marker; polymorphism; human;
KW cross-species comparison.
XX
OS Homo sapiens.
XX
PN US2003104410-A1.
XX
PD 05-JUN-2003.
XX
PF 15-MAR-2002; 2002US-00098263.
XX
PR 16-MAR-2001; 2001US-0276759P.
XX
PA (AFFY-) AFFYMETRIX INC.
XX
PI Mittmann MP;
XX
DR WPI; 2003-567953/53.
XX
PT New array of nucleic acid probes, useful for in situ hybridization, in
PT Southern, Northern or dot-blot hybridization to identify or detect the
PT sequence or specific mutations of any gene.
XX
PS Claim 1; SEQ ID NO 60525; 9pp; English.
XX
CC The invention discloses a microarray comprising a plurality of nucleic
CC acid probes including one of 2,018,500 fully defined sequences, or its
CC perfect match, perfect mismatch, antisense match or antisense mismatch.
CC Also disclosed is a method of gene expression analysis. The array is used
CC in monitoring gene expression levels by hybridisation to a DNA library,
CC in analysis of genetic variation or in hybridisation of tag-labelled
CC compounds. The nucleic acid probes are specifically designed for analysis
CC of at least one target sequence. The method of analysis comprises
CC hybridising at least one or more nucleic acids to at least two or more
CC nucleic acid probes and detecting the hybridisation. The nucleic acid
CC probes are attached to a solid support. The analysis comprises monitoring
CC gene expression levels, identifying biallelic markers or polymorphisms,
CC or family members of a gene and a cross-species comparison. Each of the
CC nucleic acids further comprises a tag sequence. The array of nucleic acid
CC probes is useful in in situ hybridisation, in Southern, Northern or dot-
CC blot hybridisation to identify or detect the sequence or specific
CC mutations of any gene, in mapping the 5' termini of mRNA molecules by
CC primer extensions or in screening cDNA or genomic libraries or subclones
CC for additional subclones containing segments of DNA that have been
CC isolated and previously sequenced. The sequence presented is one of the
CC nucleic acid probes incorporated in the microarray. Note: The sequence
CC data for this patent can also be obtained in electronic format directly
CC from USPTO at seqdata.uspto.goc/sequence.html
XX
SQ Sequence 25 BP; 7 A; 4 C; 5 G; 9 T; 0 U; 0 Other;

  Query Match 72.0%; Score 14.4; DB 8; Length 25;
  Best Local Similarity 93.8%; Pred. No. 2.8e+03;
  Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 AGTAACATCTATGTTT 16
              ||||||||| ||||||
Db          3 AGTAACATCGATGTTT 18
```

RESULT 2
AAA61982
ID AAA61982 standard; DNA; 20 BP.
XX
AC AAA61982;
XX
DT 20-NOV-2000 (first entry)
XX
DE Human MEKK5 phosphorothioate antisense oligonucleotide, SEQ ID NO:34.
XX
KW Human MEKK5; mitogen-activated protein kinase kinase kinase 5;
KW MEK kinase 5; MAP/ERK kinase kinase 5; ASK1; pro-apoptotic;
KW apoptosis signal-regulating kinase 1; programmed cell death;
KW serine/threonine kinase; MAP kinase cascade; JNK/SAPK module;
KW Jun N-terminal kinase/stress-activated protein kinase; p38 module; MKK3;
KW SEK1; transcription factor modulation; expression inhibition; antisense;

KW inflammation; wound healing disorder; phosphorothioate; ss.
XX
OS Homo sapiens.
XX
PN US6080546-A.
XX
PD 27-JUN-2000.
XX
PF 23-JUL-1999; 99US-00359757.
XX
PR 23-JUL-1999; 99US-00359757.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Monia BP, Cowsert LM, Gaarde W;
XX
DR WPI; 2000-464034/40.
XX
PT Antisense compounds useful for treating or preventing infection,
PT inflammation or tumor formation by inhibiting expression of human MEKK5.
XX
PS Claim 3; Col 39; 35pp; English.
XX
CC Sequences AAA61956-A61995 represent phosphorothioate antisense
CC oligonucleotides targetted to the human MEKK5 gene, which inhibit its
CC expression. The antisense oligonucleotides were designed to target
CC different regions of the human MEKK5 RNA, and were analysed for their
CC effect on MEKK5 mRNA levels by quantitative real-time PCR. MEKK5 (also
CC known as mitogen-activated protein kinase kinase kinase 5, MEK kinase 5,
CC MAP/ERK kinase kinase 5, apoptosis signal-regulating kinase 1, and ASK1)
CC is a dual-specific serine/threonine kinase which mediates cellular
CC responses to mitogenic stimuli by activating both the JNK/SAPK (Jun N-
CC terminal kinase/stress-activated protein kinase) and p38 modules of MAP
CC kinase cascades. MEKK5 is thought to play a critical role in the
CC regulation of apoptosis (programmed cell death) by interacting with other
CC proteins in this cascade and by phosphorylating downstream targets such
CC as MKK3 and SEK1. MEKK5 also participates in another apoptosis-related
CC signalling cascade involving the modulation of transcription factors.
CC Activation and dimerisation of MEKK5 is induced by tumour necrosis factor
CC -alpha (TNF-alpha), these processes being mediated by reactive oxygen
CC species. Thioredoxin is able to associate with MEKK5 and inhibit MEKK5
CC kinase activity and hence MEKK5-dependent apoptosis. The oligonucleotides
CC of the invention are useful for diagnosis, prevention and treatment of
CC conditions associated with MEKK5 expression, such as inflammation and
CC wound healing disorders
XX
SQ Sequence 20 BP; 4 A; 4 C; 4 G; 8 T; 0 U; 0 Other;

Query Match 69.0%; Score 13.8; DB 3; Length 20;
Best Local Similarity 88.2%; Pred. No. 5.3e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 AGTAACATCTATGTTTG 17
              ||||||||||  | ||||
Db          2 AGTAACATCTGTCTTTG 18
```

RESULT 3
AAA40172
ID AAA40172 standard; DNA; 33 BP.
XX
AC AAA40172;
XX
DT 01-NOV-2000 (first entry)
XX
DE H. pylori beta-urease-binding antibody light chain CDR1 DNA #2.
XX
KW Acid-resistant microorganism; detection; faecal; intestine; infection;
KW monoclonal antibody; light chain; complementarity determining region;
KW CDR; beta-urease; ss.
XX
OS Unidentified.
XX
PN WO200026671-A1.
XX
PD 11-MAY-2000.
XX
PF 29-OCT-1999; 99WO-EP008212.
XX
PR 29-OCT-1998; 98EP-00120517.
PR 06-NOV-1998; 98EP-00120687.
XX
PA (CONN-) CONNEX GMBH.
XX
PI Reiter C, Cullmann G, Friedrichs U, Heppner P, Lakner M;
PI Ringeis A;
XX
DR WPI; 2000-365747/31.
DR P-PSDB; AAB10016.
XX
PT Detecting infection by acid-fast microbes for diagnosis of Helicobacter
PT pylori, comprises reacting a fecal sample with two binding reagents for
PT antigens that survive intestinal passage.
XX
PS Claim 29; Page 23; 84pp; German.
XX
CC This invention describes a novel method for the detection of a mammalian
CC infection by an acid-resistant microorganism (A) by treating a faecal
CC sample with at least two different monoclonal antibodies (MAb) (or their
CC fragments or derivatives) or aptamers (collectively (I)) and detecting
CC formation of a complex (C) between (I) and the corresponding antigen of
CC (A). The first and second (I) bind to epitopes of different antigens
CC (Ag). These epitopes are present, after passage through the intestines,
CC in at least some mammals, and have either: (i) their native structure; or
CC (ii) a structure against which an antibody is produced by an animal
CC infected or immunized with (A), or its extract, lysate, derived protein
CC or fragment, or with a synthetic peptide. Practically all mammals display
CC at least one of the specified epitopes. The method is used to detect
CC infection by acid-fast bacteria, particularly of the genera Helicobacter,
CC Mycobacterium and Campylobacter, specifically H. pylori, H. hepaticus, M.
CC tuberculosis, C. jejuni and C. pylori. (I) may also be used
CC therapeutically. The method is direct and non-invasive, and provides an
CC inexpensive and easily standardizable diagnosis, despite possible
CC degradation of antigens during passage through the intestines. This
CC sequence encodes a fragment of a H. pylori beta-urease-binding antibody
CC light chain complementarity determining region CDR1 which is used to
CC illustrate the method of the invention
XX
SQ Sequence 33 BP; 10 A; 6 C; 7 G; 10 T; 0 U; 0 Other;

Query Match 69.0%; Score 13.8; DB 3; Length 33;
Best Local Similarity 88.2%; Pred. No. 5.4e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          4 AACATCTATGTTTGGTT 20
              |||||  ||||||||||
Db         13 AACATTAATGTTTGGTT 29
```

RESULT 4
AAF88123
ID AAF88123 standard; DNA; 33 BP.
XX
AC AAF88123;
XX
DT 17-JUL-2001 (first entry)
XX
DE H. pylori beta-urease derived antibody light chain CDR1 DNA #2.
XX
KW Catalase; beta-urease; antibody; antigen; detection; infection; epitope;
KW acid-resistant microorganism; complementarity determining region; CDR;
KW feces; heavy chain; light chain; ds.
XX
OS Unidentified.
XX
PN WO200127612-A2.

```
XX
PD   19-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP010057.
XX
PR   12-OCT-1999;   99EP-00120351.
PR   16-MAR-2000; 2000EP-00105592.
PR   31-MAR-2000; 2000EP-00107028.
PR   10-MAY-2000; 2000EP-00110110.
XX
PA   (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.
XX
PI   Reiter C,  Cullmann G,  Lakner M,  Truee A,  Dehnert S,  Schwartz G;
XX
DR   WPI; 2001-282086/29.
DR   P-PSDB; AAB86096.
XX
PT   Detecting infections by acid-resistant microorganisms, particularly for
PT   diagnosing Helicobacter pylori, comprises immunochromatographic detection
PT   of antigen in feces.
XX
PS   Claim 30; Page 28; 90pp; German.
XX
CC   This invention describes a novel method for detecting infection by an
CC   acid-resistant microorganism (A), in a mammal, using
CC   immunochromatography. The method is used to diagnose infection by an acid
CC   -resistant microorganism (A), in a mammal, such as Helicobacter,
CC   Campylobacter or Mycobacterium, particularly H. pylori (most preferred),
CC   H. hepatica, C. jejuni and M. tuberculosis. The method is rapid, simple,
CC   inexpensive and non-invasive, and may indicate the stage of infection. A
CC   test strip used in the method may include a filter to eliminate particles
CC   present in the sample and only a single receptor provides a reasonably
CC   secure diagnosis, with specificity and selectivity improved by detecting
CC   several epitopes (of catalase) or different antigens (catalase and beta-
CC   urease). The method can be automated. This sequence encodes a
CC   complementarity determining region (CDR) from an antibody raised against
CC   the H. pylori catalase or beta-urease antigen which is used to illustrate
CC   the method of the invention
XX
SQ   Sequence 33 BP; 10 A; 6 C; 7 G; 10 T; 0 U; 0 Other;

  Query Match             69.0%;  Score 13.8;  DB 4;  Length 33;
  Best Local Similarity   88.2%;  Pred. No. 5.4e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 AACATCTATGTTTGGTT 20
              |||||  ||||||||||
Db         13 AACATTAATGTTTGGTT 29


RESULT 5
AAF88066
ID   AAF88066 standard; DNA; 33 BP.
XX
AC   AAF88066;
XX
DT   17-JUL-2001  (first entry)
XX
DE   H. pylori beta-urease derived antibody light chain CDR1 DNA #2.
XX
KW   Heavy chain; light chain; catalase; beta-urease; detection; CDR; antigen;
KW   infection; acid-resistant microorganism; fecal; antibody; diagnosis;
KW   antibacterial; complementarity determining region; ds.
XX
OS   Unidentified.
XX
PN   WO200127613-A2.
XX
PD   19-APR-2001.
XX
PF   12-OCT-2000; 2000WO-EP010058.
XX
PR   12-OCT-1999;   99EP-00120351.
PR   16-MAR-2000; 2000EP-00105592.
PR   31-MAR-2000; 2000EP-00107028.
PR   10-MAY-2000; 2000EP-00110110.
XX
PA   (CONN-) CONNEX GES OPTIMIERUNG VON FORSCHUNG & E.
XX
PI   Reiter C,  Cullmann G,  Heppner P,  Ringeis A,  Mueller H,  Haindl E;
XX
DR   WPI; 2001-282087/29.
DR   P-PSDB; AAB86064.
XX
PT   Detecting infections by acid-resistant microorganisms, particularly for
PT   diagnosing Helicobacter pylori, comprises an immunoassay on a fecal
PT   sample.
XX
PS   Claim 26; Page 18; 89pp; German.
XX
CC   This invention describes a novel method for detecting, in a mammal,
CC   infection by an acid-resistant microorganism (A) which comprises reacting
CC   a fecal sample with: (i) a receptor (R) such that a complex is formed
CC   with an antigen (Ag) of (A); or (ii) two different R so that a three-part
CC   complex is formed with Ag, and the formation of a complex detected. R are
CC   specific for an Ag which, after passage through the intestines, at least
CC   in some mammals, retains a native (or corresponding) structure against
CC   which the mammal produces antibodies (when immunized or infected with
CC   (A), or its extracts, lysates or derived proteins (or fragments) or
CC   synthetic peptides). The products of the invention have antibacterial
CC   activity. The method is used to diagnose infection by Helicobacter,
CC   Campylobacter or Mycobacterium, particularly H. pylori (most preferred),
CC   H. hepatica, C. jejuni and M. tuberculosis, and also to monitor the
CC   progress of treatment. Receptors, particularly antibodies, directed
CC   against Ag can be used therapeutically for treatment of infections. The
CC   method requires only one R to provide a reasonably secure diagnosis
CC   (although use of two R improves sensitivity), so is relatively
CC   inexpensive and more easily standardized. Also it is direct, non-
CC   invasive, suitable for automation and may indicate the stage of an
CC   infection. This sequence encodes a complementarity determining region
CC   (CDR) from an antibody generated against a Helicobacter pylori antigen
CC   (catalase or beta-urease) which is used to illustrate the method of the
CC   invention
XX
SQ   Sequence 33 BP; 10 A; 6 C; 7 G; 10 T; 0 U; 0 Other;

  Query Match             69.0%;  Score 13.8;  DB 4;  Length 33;
  Best Local Similarity   88.2%;  Pred. No. 5.4e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 AACATCTATGTTTGGTT 20
              |||||  ||||||||||
Db         13 AACATTAATGTTTGGTT 29


RESULT 6
AAA13992
ID   AAA13992 standard; DNA; 43 BP.
XX
AC   AAA13992;
XX
DT   08-AUG-2000  (first entry)
XX
DE   Geranylgeranyl diphosphate synthase PCR primer SEQ ID NO:9.
XX
KW   Geranylgeranyl diphosphate synthase; GGPP synthase; yew; cytostatic;
KW   anticancer; Taxus; diterpene; paclitaxel; identification; plant;
KW   Taxomyces andreanae; Penicillium raistrickii; microorganism; PCR primer;
KW   ss.
XX
OS   Taxus canadensis.
XX
PN   US6043072-A.
XX
PD   28-MAR-2000.
XX
```

PF 05-NOV-1998; 98US-00187050.
XX
PR 05-NOV-1998; 98US-00187050.
XX
PA (UNIW) UNIV WASHINGTON STATE RES FOUND.
XX
PI Croteau RB, Hefner JL;
XX
DR WPI; 2000-282526/24.
XX
PT Nucleic acid encoding geranylgeranyl diphosphate is useful for producing
PT paclitaxel and other diterpenes that are useful as anticancer drugs.
XX
PS Example 2; Col 41; 57pp; English.
XX
CC The present sequence represents a PCR primer for a geranylgeranyl
CC diphosphate (GGPP) synthase protein. GGPP synthase has cytostatic
CC activity. A vector encoding GGPP synthase is useful in increasing GGPP
CC synthase levels in a host cell preferably Taxus (Yew) cell and thereby
CC facilitates production, isolation and purification of larger amounts of
CC GGPP synthase in plants. GGPP synthase is useful in obtaining expression
CC or enhanced expression of GGPP and other diterpenes, such as paclitaxel,
CC useful as anticancer drugs. Isolated nucleic acids encoding GGPP synthase
CC or hybridising with GGPP synthase encoding nucleic acids are used for
CC identifying genes encoding GGPP synthase from microorganisms such as
CC Taxomyces andreanae and Penicillium raistrickii
XX
SQ Sequence 43 BP; 14 A; 6 C; 10 G; 13 T; 0 U; 0 Other;

Query Match 69.0%; Score 13.8; DB 3; Length 43;
Best Local Similarity 88.2%; Pred. No. 5.5e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 AACATCTATGTTTGGTT 20
|| |||||||||| ||
Db 2 AAGATCTATGTTTGATT 18

RESULT 7
ABL99963/c
ID ABL99963 standard; DNA; 24 BP.
XX
AC ABL99963;
XX
DT 08-AUG-2002 (first entry)
XX
DE HOMO 2-hydroxy acid dehydrogenase family protein 26.29 PCR primer 2.
XX
KW Human; DNA dependent 2-hydroxy acid dehydrogenase protein; enzyme;
KW cytostatic; virucidal; immunomodulatory; antiinflammatory; haemostatic;
KW malignant tumour; human immunodeficiency virus; HIV; infection;
KW immunological disease; gene therapy; PCR; primer; ss.
XX
OS Homo sapiens.
XX
PN WO200232950-A1.
XX
PD 25-APR-2002.
XX
PF 02-JUL-2001; 2001WO-CN001128.
XX
PR 07-JUL-2000; 2000CN-00117047.
XX
PA (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2002-340232/37.
XX
PT Polypeptide-HOMO DNA-dependent 2-hydroxy acid dehydrogenase family
PT protein 26.29 and encoding polynucleotide, used in diagnosis and
PT treatment of e.g. malignant tumors, hemopathy, immunological diseases and
PT phlogosis.
XX
PS Example 2; Page 13; 37pp; Chinese.
XX
CC The invention relates to HOMO DNA-dependent 2-hydroxy acid dehydrogenase
CC family protein 26.29 with cytostatic, virucidal, immunomodulatory,
CC antiinflammatory and haemostatic activity. The protein and encoding
CC polynucleotide are used in diagnosis and treatment of malignant tumour,
CC haemopathy, human immunodeficiency virus (HIV) infection, immunological
CC diseases and various inflammations. The polynucleotide is useful in gene
CC therapy. The present sequence is that of a HOMO DNA-dependent 2-hydroxy
CC acid dehydrogenase family protein 26.29 PCR primer, useful in examples of
CC the invention
XX
SQ Sequence 24 BP; 13 A; 2 C; 2 G; 7 T; 0 U; 0 Other;

Query Match 68.0%; Score 13.6; DB 6; Length 24;
Best Local Similarity 80.0%; Pred. No. 6.6e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AGTAACATCTATGTTTGGTT 20
|||||| ||||| ||| | |
Db 22 AGTAACTTCTATATTTTGAT 3

RESULT 8
AAF77839/c
ID AAF77839 standard; DNA; 28 BP.
XX
AC AAF77839;
XX
DT 04-JUN-2001 (first entry)
XX
DE Glycerol dehydratase promoter sequence.
XX
KW Glycerol dehydratase; 1,3-propanediol; enzyme stabiliser;
KW polymer production; promoter; enzyme; ss.
XX
OS Unidentified.
XX
PN FR2796081-A1.
XX
PD 12-JAN-2001.
XX
PF 09-JUL-1999; 99FR-00008939.
XX
PR 09-JUL-1999; 99FR-00008939.
XX
PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
PA (NASC-) INST NAT SCI APPLIQUEES TOULOUSE.
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Sarcabal P, Croux C, Soucaille P;
XX
DR WPI; 2001-247136/26.
XX
PT Production of 1,3-propanediol comprises culturing a recombinant
PT microorganism expressing coenzyme B12-independent glycerol dehydratase.
XX
PS Claim 18; Page 37; 69pp; French.
XX
CC The present invention relates to a method for producing 1,3-propanediol
CC from a carbon source, comprising culturing a recombinant microorganism
CC having glycerol dehydratase and/or 1,3-propanediol dehydrogenase from
CC Clostridium butyricum (see AAB73300, AAB80887 and AAB80888). The method
CC of the present invention is useful for the production of 1,3-propanediol,
CC which is useful as a stabiliser for lipases, amylases and proteases in
CC wash liquids, as a protective emollient in liquid detergents for hand and
CC dish washing, and as a monomer for producing polymers, especially
CC polyesters, polyethers and polyurethanes. The present sequence is a
CC promoter sequence, for glycerol dehydratase, which was used in the
CC present invention
XX
SQ Sequence 28 BP; 16 A; 2 C; 2 G; 8 T; 0 U; 0 Other;

```
  Query Match             68.0%;  Score 13.6;  DB 4;  Length 28;
  Best Local Similarity   80.0%;  Pred. No. 6.7e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGTT 20
              | |||||| | |||||| ||
Db         26 AATAACATTTTTGTTTGTTT 7


RESULT 9
ACI17115/c
ID   ACI17115 standard; DNA; 25 BP.
XX
AC   ACI17115;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human microarray DNA oligonucleotide SEQ ID NO 17106.
XX
KW   EST; ss; probe; expressed sequence tag; microarray; gene expression;
KW   genetic variation; biallelic marker; polymorphism; human;
KW   cross-species comparison.
XX
OS   Homo sapiens.
XX
PN   US2003104410-A1.
XX
PD   05-JUN-2003.
XX
PF   15-MAR-2002; 2002US-00098263.
XX
PR   16-MAR-2001; 2001US-0276759P.
XX
PA   (AFFY-) AFFYMETRIX INC.
XX
PI   Mittmann MP;
XX
DR   WPI; 2003-567953/53.
XX
PT   New array of nucleic acid probes, useful for in situ hybridization, in
PT   Southern, Northern or dot-blot hybridization to identify or detect the
PT   sequence or specific mutations of any gene.
XX
PS   Claim 1; SEQ ID NO 17106; 9pp; English.
XX
CC   The invention discloses a microarray comprising a plurality of nucleic
CC   acid probes including one of 2,018,500 fully defined sequences, or its
CC   perfect match, perfect mismatch, antisense match or antisense mismatch.
CC   Also disclosed is a method of gene expression analysis. The array is used
CC   in monitoring gene expression levels by hybridisation to a DNA library,
CC   in analysis of genetic variation or in hybridisation of tag-labelled
CC   compounds. The nucleic acid probes are specifically designed for analysis
CC   of at least one target sequence. The method of analysis comprises
CC   hybridising at least one or more nucleic acids to at least two or more
CC   nucleic acid probes and detecting the hybridisation. The nucleic acid
CC   probes are attached to a solid support. The analysis comprises monitoring
CC   gene expression levels, identifying biallelic markers or polymorphisms,
CC   or family members of a gene and a cross-species comparison. Each of the
CC   nucleic acids further comprises a tag sequence. The array of nucleic acid
CC   probes is useful in in situ hybridisation, in Southern, Northern or dot-
CC   blot hybridisation to identify or detect the sequence or specific
CC   mutations of any gene, in mapping the 5' termini of mRNA molecules by
CC   primer extensions or in screening cDNA or genomic libraries or subclones
CC   for additional subclones containing segments of DNA that have been
CC   isolated and previously sequenced. The sequence presented is one of the
CC   nucleic acid probes incorporated in the microarray. Note: The sequence
CC   data for this patent can also be obtained in electronic format directly
CC   from USPTO at seqdata.uspto.goc/sequence.html
XX
SQ   Sequence 25 BP; 9 A; 5 C; 8 G; 3 T; 0 U; 0 Other;

  Query Match             67.0%;  Score 13.4;  DB 8;  Length 25;
  Best Local Similarity   93.3%;  Pred. No. 8.3e+03;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 CATCTATGTTTGGTT 20
              ||||| |||||||||
Db         16 CATCTCTGTTTGGTT 2


RESULT 10
AAX06450/c
ID   AAX06450 standard; DNA; 31 BP.
XX
AC   AAX06450;
XX
DT   31-MAR-1999  (first entry)
XX
DE   Human biallelic polymorphic DNA fragment SGC34498.
XX
KW   Polymorphism; biallelic; paternity testing; forensic; genetic mapping;
KW   phenotypic typing; medicament; disease; marker; human; ss.
XX
OS   Homo sapiens.
XX
PN   WO9858529-A2.
XX
PD   30-DEC-1998.
XX
PF   22-JUN-1998;   98WO-US012930.
XX
PR   24-JUN-1997;   97US-0050594P.
XX
PA   (AFFY-) AFFYMETRIX INC.
XX
PI   Lipshutz RJ,  Chee M,  Fan J,  Berno A;
XX
DR   WPI; 1999-080963/07.
XX
PT   New nucleic acid segments containing polymorphic sites - used for, e.g.
PT   detecting a disease phenotype, in forensics, paternity testing or genetic
PT   mapping of phenotypic traits.
XX
PS   Claim 1; Page 28; 61pp; English.
XX
CC   Sequences AAX06101-X06558 represent human DNA fragments which contain
CC   biallelic polymorphic markers. The base occupying the polymorphic site is
CC   indicated by the appropriate IUPAC-IUB ambiguity code. These fragments
CC   can be used in a method for determining polymorphic forms in an
CC   individual. The invention further provides computer-readable storage
CC   medium for storing data for access by an application programme being
CC   executed on a data processing system. Such a method comprises a data
CC   structure stored in the computer-readable storage medium, the data
CC   structure including information resident in a database used by the
CC   application programme and including records, each record comprising
CC   information identifying a polymorphism shown in the above sequences. The
CC   products and methods can be used for analysing polymorphic sites in
CC   individuals for testing for the presence of a disease phenotype or in
CC   forensics, paternity testing or genetic mapping of phenotypic traits.
CC   They can also be used for the production of polypeptides expressed by
CC   variant genes and for the production of transgenic animals. The nucleic
CC   acid segments can also be used in the manufacture of medicaments for the
CC   treatment or prophylaxis of diseases
XX
SQ   Sequence 31 BP; 13 A; 6 C; 6 G; 5 T; 0 U; 1 Other;

  Query Match             67.0%;  Score 13.4;  DB 2;  Length 31;
  Best Local Similarity   82.4%;  Pred. No. 8.4e+03;
  Matches   14;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 AACATCTATGTTTGGTT 20
              |||| || |||||:|||
Db         29 AACACCTTTGTTTKGTT 13
```

RESULT 11
AAT89524
ID AAT89524 standard; cDNA; 24 BP.
XX
AC AAT89524;
XX
DT 27-JAN-1998 (first entry)
XX
DE Bloom's syndrome active BLM gene SSCP forward primer C1-3.
XX
KW BLM; Bloom's syndrome; BS; mutant; probe; PCR primer; cancer; therapy;
KW diagnosis; SSCP; Single-Strand Conformation Polymorphism; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9717979-A1.
XX
PD 22-MAY-1997.
XX
PF 15-NOV-1996; 96WO-US019046.
XX
PR 15-NOV-1995; 95US-00559303.
XX
PA (NYBL-) NEW YORK BLOOD CENT INC.
XX
PI Ellis N, German J, Groden J;
XX
DR WPI; 1997-289051/26.
XX
PT Diagnosing Bloom's syndrome, and carriers, by detecting mutant BLM genes
PT - for gene therapy with nucleic acid encoding active BLM protein to treat
PT Bloom's syndrome and cancer in general.
XX
PS Disclosure; Page 31; 51pp; English.
XX
CC This forward primer is used in the PCR amplification of the BLM gene
CC sequence that encodes an enzymatically active BLM protein. This is used
CC in the Single-Strand Conformation Polymorphism (SSCP) Analysis of the BLM
CC gene. SSCP analysis helps in identifying the mutants in the BLM gene.
CC Bloom's syndrome is diagnosed by detecting 2 mutated BLM genes or the
CC absence of a wild-type BLM gene in a subject. Delivery of a functional
CC BLM gene to bone marrow cells is used to treat or prevent the onset of
CC Bloom's syndrome. Identification of the BLM gene and its products should
CC assist in the development of therapeutic and diagnostic agents for cancer
XX
SQ Sequence 24 BP; 7 A; 4 C; 5 G; 8 T; 0 U; 0 Other;

Query Match 66.0%; Score 13.2; DB 2; Length 24;
Best Local Similarity 83.3%; Pred. No. 1e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 AGTAACATCTATGTTTGG 18
              |||| |||| ||| ||||
Db          6 AGTACCATCAATGATTGG 23
```

RESULT 12
ABZ29491
ID ABZ29491 standard; DNA; 65 BP.
XX
AC ABZ29491;
XX
DT 30-JAN-2003 (first entry)
XX
DE Candida gene related tetracyclin promoter PCR primer SEQ ID NO 3574.
XX
KW Fungus; yeast; tetracyclin; promoter; GRACE strain; biosynthesis;
KW signal transduction; DNA replication; cell division; growth;
KW proliferation; Candida albicans; fungicide; antifungal; PCR; primer; ss.
XX
OS Candida albicans.
XX
PN WO200253728-A2.
XX
PD 11-JUL-2002.
XX
PF 26-DEC-2001; 2001WO-US049486.
XX
PR 29-DEC-2000; 2000US-0259128P.
PR 20-FEB-2001; 2001US-00792024.
PR 22-AUG-2001; 2001US-0314050P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;
XX
DR WPI; 2002-566694/60.
XX
PT Constructing strains for identifying gene products as effective targets
PT for therapeutic intervention, by inactivating in the strain one allele of
PT a gene and placing other allele of the gene under conditional expression.
XX
PS Claim 76; SEQ ID NO 3574; 167pp + Sequence Listing; English.
XX
CC The invention relates to constructing (M1) a strain of diploid fungal
CC cells in which both alleles of a gene are modified, comprising modifying
CC one allele by insertion or replacement by a cassette having an
CC expressible selectable marker and modifying other allele by
CC recombination, of a promoter replacement fragment with a heterologous
CC promoter, so that expression of the second allele is regulated by the
CC promoter. (M1) is useful for constructing a strain of diploid fungal
CC cells in which both alleles of a gene are modified. The diploid fungal
CC cells having both alleles modified are useful for identifying a gene that
CC is essential to the survival or growth of a fungus, a gene that
CC contributes to the virulence and/or pathogenicity of a fungus, a gene
CC that contributes to the resistance of a diploid fungus to an antifungal
CC agent, an antifungal agent that inhibits the growth of a diploid fungus
CC and for identifying a therapeutic agent for treatment of a mammalian
CC disease. (M1) is useful for identifying a compound which modulates the
CC activity of a gene product, preferably enzymatic activity, carbon
CC compound catabolism, biosynthetic, transporter, transcriptional,
CC translational, signal transduction, DNA replication and cell division
CC activity. The method is useful for identifying a compound having the
CC ability to inhibit growth or proliferation of C. albicans cells and for
CC treating infection by C. albicans. The present sequence is that of a PCR
CC primer used in the method of the invention. Note: The sequence data for
CC this patent is not represented in the printed specification but is based
CC on sequence information supplied to Derwent by the European Patent Office
XX
SQ Sequence 65 BP; 17 A; 12 C; 14 G; 22 T; 0 U; 0 Other;

Query Match 66.0%; Score 13.2; DB 6; Length 65;
Best Local Similarity 83.3%; Pred. No. 1.1e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          2 GTAACATCTATGTTTGGT 19
              ||||||| | |||||||
Db          7 GTAACATTCAAGTTTGGT 24
```

RESULT 13
ACD50544
ID ACD50544 standard; RNA; 17 BP.
XX
AC ACD50544;
XX
DT 23-SEP-2003 (first entry)
XX
DE HBV hammerhead ribozyme substrate sequence #112.
XX
KW Nucleic acid molecule; Hepatitis C virus; HCV; Hepatitis B virus; HBV;
KW RNA stability; RNA expression; RNA synthesis; antisense;
KW enzymatic nucleic acid; hammerhead ribozyme; DNAzyme; inozyme; zinzyme;
KW amberzyme; G-cleaver ribozyme; decoy molecule; aptamer;
KW HBV reverse transcriptase; Enhancer I region; viral replication;

KW degenerative; disease state; HBV infection; HCV infection; cirrhosis;
KW liver failure; hepatocellular carcinoma; hepatotropic; cytostatic;
KW virucide; antiinflammatory; substrate; ss.
XX
OS Hepatitis B virus.
XX
PN WO200281494-A1.
XX
PD 17-OCT-2002.
XX
PF 26-MAR-2002; 2002WO-US009187.
XX
PR 26-MAR-2001; 2001US-00817879.
PR 08-JUN-2001; 2001US-00877478.
PR 08-JUN-2001; 2001US-0296876P.
PR 24-OCT-2001; 2001US-0335059P.
PR 05-DEC-2001; 2001US-0337055P.
XX
PA (RIBO-) RIBOZYME PHARM INC.
PA (BLAT/) BLATT L.
PA (MACE/) MACEJAK D.
PA (MCSW/) MCSWIGGEN J.
PA (MORR/) MORRISSEY D.
PA (PAVC/) PAVCO P.
PA (LEEP/) LEE P.
PA (DRAP/) DRAPER K.
PA (ROBE/) ROBERTS E.
XX
PI Blatt L, Macejak D, Mcswiggen J, Morrissey D, Pavco P, Lee P;
PI Draper K, Roberts E;
XX
DR *WPI; 2003-229207/22.*
XX
PT Novel compound useful for treating cirrhosis, liver failure,
PT hepatocellular carcinoma, or condition associated with hepatitis C virus
PT infection.
XX
PS Example 1; Page 138; 387pp; English.
XX
CC The present invention relates to nucleic acid molecules which modulate
CC the synthesis, expression and/or stability of Hepatitis C virus (HCV) or
CC Hepatitis B virus (HBV) RNA. The nucleic acid molecules include antisense
CC and enzymatic nucleic acids such as hammerhead ribozymes, DNAzymes,
CC inozymes, zinzymes, amberzymes, and G-cleaver ribozymes. Also disclosed
CC are nucleic acid decoy molecules and aptamers that bind to HBV reverse
CC transcriptase and/or HBV reverse transcriptase primer sequences, as well
CC as oligonucleotides that specifically bind the Enhancer I region of HBV
CC DNA. The nucleic acids may be used to modulate the expression of HBV
CC genes and HBV viral replication. Also disclosed is a method for screening
CC compounds and/or potential therapies directed against HBV, and compounds
CC that modulate the expression and/or replication of HCV. The compounds and
CC methods of the invention are useful for the treatment of degenerative and
CC disease states related to HBV and HCV infection, replication and gene
CC expression such as cirrhosis, liver failure, and hepatocellular
CC carcinoma. The present sequence represents a substrate for one of the HBV
CC ribozyme, inozyme, G-cleaver, zinzyme, DNAzyme or amberzyme sequences
CC disclosed in the present invention
XX
SQ Sequence 17 BP; 4 A; 4 C; 3 G; 0 T; 6 U; 0 Other;

Query Match 64.0%; Score 12.8; DB 7; Length 17;
Best Local Similarity 50.0%; Pred. No. 1.6e+04;
Matches 8; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 AGTAACATCTATGTTT 16
              || ||| :|:|:|:::
Db          1 AGGAACCUCUAUGUUU 16
```

RESULT 14
ACD51915
ID ACD51915 standard; RNA; 17 BP.
XX
AC ACD51915;
XX
DT 24-SEP-2003 (first entry)
XX
DE HBV inozyme substrate sequence #147.
XX
KW Nucleic acid molecule; Hepatitis C virus; HCV; Hepatitis B virus; HBV;
KW RNA stability; RNA expression; RNA synthesis; antisense;
KW enzymatic nucleic acid; hammerhead ribozyme; DNAzyme; inozyme; zinzyme;
KW amberzyme; G-cleaver ribozyme; decoy molecule; aptamer;
KW HBV reverse transcriptase; Enhancer I region; viral replication;
KW degenerative; disease state; HBV infection; HCV infection; cirrhosis;
KW liver failure; hepatocellular carcinoma; hepatotropic; cytostatic;
KW virucide; antiinflammatory; substrate; ss.
XX
OS Hepatitis B virus.
XX
PN WO200281494-A1.
XX
PD 17-OCT-2002.
XX
PF 26-MAR-2002; 2002WO-US009187.
XX
PR 26-MAR-2001; 2001US-00817879.
PR 08-JUN-2001; 2001US-00877478.
PR 08-JUN-2001; 2001US-0296876P.
PR 24-OCT-2001; 2001US-0335059P.
PR 05-DEC-2001; 2001US-0337055P.
XX
PA (RIBO-) RIBOZYME PHARM INC.
PA (BLAT/) BLATT L.
PA (MACE/) MACEJAK D.
PA (MCSW/) MCSWIGGEN J.
PA (MORR/) MORRISSEY D.
PA (PAVC/) PAVCO P.
PA (LEEP/) LEE P.
PA (DRAP/) DRAPER K.
PA (ROBE/) ROBERTS E.
XX
PI Blatt L, Macejak D, Mcswiggen J, Morrissey D, Pavco P, Lee P;
PI Draper K, Roberts E;
XX
DR WPI; 2003-229207/22.
XX
PT Novel compound useful for treating cirrhosis, liver failure,
PT hepatocellular carcinoma, or condition associated with hepatitis C virus
PT infection.
XX
PS Example 1; Page 152; 387pp; English.
XX
CC The present invention relates to nucleic acid molecules which modulate
CC the synthesis, expression and/or stability of Hepatitis C virus (HCV) or
CC Hepatitis B virus (HBV) RNA. The nucleic acid molecules include antisense
CC and enzymatic nucleic acids such as hammerhead ribozymes, DNAzymes,
CC inozymes, zinzymes, amberzymes, and G-cleaver ribozymes. Also disclosed
CC are nucleic acid decoy molecules and aptamers that bind to HBV reverse
CC transcriptase and/or HBV reverse transcriptase primer sequences, as well
CC as oligonucleotides that specifically bind the Enhancer I region of HBV
CC DNA. The nucleic acids may be used to modulate the expression of HBV
CC genes and HBV viral replication. Also disclosed is a method for screening
CC compounds and/or potential therapies directed against HBV, and compounds
CC that modulate the expression and/or replication of HCV. The compounds and
CC methods of the invention are useful for the treatment of degenerative and
CC disease states related to HBV and HCV infection, replication and gene
CC expression such as cirrhosis, liver failure, and hepatocellular
CC carcinoma. The present sequence represents a substrate for one of the HBV
CC ribozyme, inozyme, G-cleaver, zinzyme, DNAzyme or amberzyme sequences
CC disclosed in the present invention
XX
SQ Sequence 17 BP; 5 A; 3 C; 3 G; 0 T; 6 U; 0 Other;

Query Match 64.0%; Score 12.8; DB 7; Length 17;
Best Local Similarity 50.0%; Pred. No. 1.6e+04;

```
  Matches    8;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTT 16
              || ||| :|:|:|::: 
Db          2 AGGAACCUCUAUGUUU 17


RESULT 15
AAQ21867/c
ID   AAQ21867 standard; DNA; 25 BP.
XX
AC   AAQ21867;
XX
DT   23-JUN-1992  (first entry)
XX
DE   Probe 179.3 DNA, to demonstrate controlled amplification.
XX
KW   Templates; Human papilloma virus; ligase chain reaction; LCR; ss.
XX
OS   Synthetic.
XX
PN   EP473155-A.
XX
PD   04-MAR-1992.
XX
PF   29-AUG-1991;   91EP-00114541.
XX
PR   30-AUG-1990;   90US-00575177.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Backman KC,  Carrino JJ,  Shimer GH;
XX
DR   WPI; 1992-073668/10.
XX
PT   Target-dependent prodn. of templates for ligase chain reaction -
PT   increases sensitivity and detection of target from non-target contg.
PT   samples.
XX
PS   Example 4; Page 7; 20pp; English.
XX
CC   Duplex DNA (AAQ21864) (adapted from the L1 region of Human papilloma
CC   virus type 16 ) was used as a target to study the linear pre-
CC   amplification process. A probe set (QQ21865-8) was designed to hybridise
CC   to the target sequence for use in pre- amplification. The 3' end of probe
CC   179.3 was haptenated with biotin-aminocaproyl NHS active ester. See also
CC   AAQ21862-78
XX
SQ   Sequence 25 BP; 11 A; 6 C; 2 G; 6 T; 0 U; 0 Other;

  Query Match             64.0%;  Score 12.8;  DB 2;  Length 25;
  Best Local Similarity   87.5%;  Pred. No. 1.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTG 17
              |||| || ||||||||
Db         24 GTAATATATATGTTTG 9


RESULT 16
ACI60535
ID   ACI60535 standard; DNA; 25 BP.
XX
AC   ACI60535;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human microarray DNA oligonucleotide SEQ ID NO 60526.
XX
KW   EST; ss; probe; expressed sequence tag; microarray; gene expression;
KW   genetic variation; biallelic marker; polymorphism; human;
KW   cross-species comparison.
XX
OS   Homo sapiens.
XX
PN   US2003104410-A1.
XX
PD   05-JUN-2003.
XX
PF   15-MAR-2002; 2002US-00098263.
XX
PR   16-MAR-2001; 2001US-0276759P.
XX
PA   (AFFY-) AFFYMETRIX INC.
XX
PI   Mittmann MP;
XX
DR   WPI; 2003-567953/53.
XX
PT   New array of nucleic acid probes, useful for in situ hybridization, in
PT   Southern, Northern or dot-blot hybridization to identify or detect the
PT   sequence or specific mutations of any gene.
XX
PS   Claim 1; SEQ ID NO 60526; 9pp; English.
XX
CC   The invention discloses a microarray comprising a plurality of nucleic
CC   acid probes including one of 2,018,500 fully defined sequences, or its
CC   perfect match, perfect mismatch, antisense match or antisense mismatch.
CC   Also disclosed is a method of gene expression analysis. The array is used
CC   in monitoring gene expression levels by hybridisation to a DNA library,
CC   in analysis of genetic variation or in hybridisation of tag-labelled
CC   compounds. The nucleic acid probes are specifically designed for analysis
CC   of at least one target sequence. The method of analysis comprises
CC   hybridising at least one or more nucleic acids to at least two or more
CC   nucleic acid probes and detecting the hybridisation. The nucleic acid
CC   probes are attached to a solid support. The analysis comprises monitoring
CC   gene expression levels, identifying biallelic markers or polymorphisms,
CC   or family members of a gene and a cross-species comparison. Each of the
CC   nucleic acids further comprises a tag sequence. The array of nucleic acid
CC   probes is useful in in situ hybridisation, in Southern, Northern or dot-
CC   blot hybridisation to identify or detect the sequence or specific
CC   mutations of any gene, in mapping the 5' termini of mRNA molecules by
CC   primer extensions or in screening cDNA or genomic libraries or subclones
CC   for additional subclones containing segments of DNA that have been
CC   isolated and previously sequenced. The sequence presented is one of the
CC   nucleic acid probes incorporated in the microarray. Note: The sequence
CC   data for this patent can also be obtained in electronic format directly
CC   from USPTO at seqdata.uspto.goc/sequence.html
XX
SQ   Sequence 25 BP; 6 A; 4 C; 5 G; 10 T; 0 U; 0 Other;

  Query Match             64.0%;  Score 12.8;  DB 8;  Length 25;
  Best Local Similarity   87.5%;  Pred. No. 1.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTT 16
              |||||||||  |||||
Db          3 AGTAACATCGTTGTTT 18


RESULT 17
ACI01532
ID   ACI01532 standard; DNA; 25 BP.
XX
AC   ACI01532;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human microarray DNA oligonucleotide SEQ ID NO 1523.
XX
KW   EST; ss; probe; expressed sequence tag; microarray; gene expression;
KW   genetic variation; biallelic marker; polymorphism; human;
KW   cross-species comparison.
XX
OS   Homo sapiens.
XX
```

PN US2003104410-A1.
XX
PD 05-JUN-2003.
XX
PF 15-MAR-2002; 2002US-00098263.
XX
PR 16-MAR-2001; 2001US-0276759P.
XX
PA (AFFY-) AFFYMETRIX INC.
XX
PI Mittmann MP;
XX
DR WPI; 2003-567953/53.
XX
PT New array of nucleic acid probes, useful for in situ hybridization, in
PT Southern, Northern or dot-blot hybridization to identify or detect the
PT sequence or specific mutations of any gene.
XX
PS Claim 1; SEQ ID NO 1523; 9pp; English.
XX
CC The invention discloses a microarray comprising a plurality of nucleic
CC acid probes including one of 2,018,500 fully defined sequences, or its
CC perfect match, perfect mismatch, antisense match or antisense mismatch.
CC Also disclosed is a method of gene expression analysis. The array is used
CC in monitoring gene expression levels by hybridisation to a DNA library,
CC in analysis of genetic variation or in hybridisation of tag-labelled
CC compounds. The nucleic acid probes are specifically designed for analysis
CC of at least one target sequence. The method of analysis comprises
CC hybridising at least one or more nucleic acids to at least two or more
CC nucleic acid probes and detecting the hybridisation. The nucleic acid
CC probes are attached to a solid support. The analysis comprises monitoring
CC gene expression levels, identifying biallelic markers or polymorphisms,
CC or family members of a gene and a cross-species comparison. Each of the
CC nucleic acids further comprises a tag sequence. The array of nucleic acid
CC probes is useful in in situ hybridisation, in Southern, Northern or dot-
CC blot hybridisation to identify or detect the sequence or specific
CC mutations of any gene, in mapping the 5' termini of mRNA molecules by
CC primer extensions or in screening cDNA or genomic libraries or subclones
CC for additional subclones containing segments of DNA that have been
CC isolated and previously sequenced. The sequence presented is one of the
CC nucleic acid probes incorporated in the microarray. Note: The sequence
CC data for this patent can also be obtained in electronic format directly
CC from USPTO at seqdata.uspto.goc/sequence.html
XX
SQ Sequence 25 BP; 4 A; 4 C; 5 G; 12 T; 0 U; 0 Other;

Query Match 64.0%; Score 12.8; DB 8; Length 25;
Best Local Similarity 87.5%; Pred. No. 1.6e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          4 AACATCTATGTTTGGT 19
              ||| ||||| ||||||
Db          1 AACGTCTATCTTTGGT 16
```

RESULT 18
ACH61845
ID ACH61845 standard; DNA; 25 BP.
XX
AC ACH61845;
XX
DT 17-OCT-2003 (first entry)
XX
DE DNA target sequence #10981 useful in array for genetic analyses.
XX
KW Gene expression analysis; array; hybridisation; genetic variation;
KW tag-labelled compound; gene family; in situ hybridisation;
KW library screening; Southern hybridisation; northern hybridisation;
KW dot-blot hybridisation; gene sequence; mutation detection;
KW target sequence; probe; PCR; primer; ss.
XX
OS Unidentified.
XX

PN US2003082596-A1.
XX
PD 01-MAY-2003.
XX
PF 08-AUG-2002; 2002US-00215112.
XX
PR 08-AUG-2001; 2001US-0311040P.
XX
PA (MITT/) MITTMANN M.
XX
PI Mittmann M;
XX
DR WPI; 2003-576608/54.
XX
PT New probe array useful e.g. for monitoring gene expression levels, for
PT analyzing genetic variations, or for hybridizing tag-labeled compounds,
PT comprises multiple nucleic acid probes.
XX
PS Claim 1; SEQ ID NO 10981; 9pp; English.
XX
CC The present invention relates to nucleic acid sequences that are
CC complementary to particular genes, and can be used as probes for a
CC variety of analyses such as gene expression analysis. Each probe
CC comprises 9 or more consecutive nucleotides from at least one of 14936
CC nucleotide sequences defined in the patent, or their perfect sense match,
CC sense mismatch, antisense match or antisense mismatch oligonucleotides.
CC The probes may be used in an array comprising at least 10 distinct
CC nucleic acid probes. The array is useful in monitoring gene expression
CC levels by hybridisation to a DNA library, in analysing genetic
CC variations, and in hybridising tag-labelled compounds. The probes are
CC useful for identifying family members of a gene. The probes are also
CC useful in in situ hybridisations, in screening cDNA or genomic libraries
CC (or derived subclones) for additional clones containing segments of DNA
CC that have been previously isolated and sequenced, in Southern, northern,
CC or dot-blot hybridisation of genomic DNA to identify or detect the
CC sequence of any gene or detect specific mutations in any gene, and in
CC mapping the 5' termini of mRNA molecules by primer extensions. The
CC nucleic acid sequences of the invention are also useful as PCR primers.
CC The invention provides a large collection of nucleic acid sequences
CC complementary to particular genes with a wide range of analytical uses.
CC ACH50865-ACH65260 represent the target sequences of the invention. Note:
CC The sequence data for this patent was obtained in electronic format
CC directly from the USPTO web site at seqdata.uspto.gov/psipsDIDEntry.html
XX
SQ Sequence 25 BP; 5 A; 4 C; 4 G; 12 T; 0 U; 0 Other;

Query Match 64.0%; Score 12.8; DB 8; Length 25;
Best Local Similarity 87.5%; Pred. No. 1.6e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          4 AACATCTATGTTTGGT 19
              ||| |||||||||| |
Db          4 AACTTCTATGTTTGCT 19
```

RESULT 19
ACH57384
ID ACH57384 standard; DNA; 25 BP.
XX
AC ACH57384;
XX
DT 16-OCT-2003 (first entry)
XX
DE DNA target sequence #6520 useful in array for genetic analyses.
XX
KW Gene expression analysis; array; hybridisation; genetic variation;
KW tag-labelled compound; gene family; in situ hybridisation;
KW library screening; Southern hybridisation; northern hybridisation;
KW dot-blot hybridisation; gene sequence; mutation detection;
KW target sequence; probe; PCR; primer; ss.
XX
OS Unidentified.
XX

PN US2003082596-A1.
XX
PD 01-MAY-2003.
XX
PF 08-AUG-2002; 2002US-00215112.
XX
PR 08-AUG-2001; 2001US-0311040P.
XX
PA (MITT/) MITTMANN M.
XX
PI Mittmann M;
XX
DR WPI; 2003-576608/54.
XX
PT New probe array useful e.g. for monitoring gene expression levels, for
PT analyzing genetic variations, or for hybridizing tag-labeled compounds,
PT comprises multiple nucleic acid probes.
XX
PS Claim 1; SEQ ID NO 6520; 9pp; English.
XX
CC The present invention relates to nucleic acid sequences that are
CC complementary to particular genes, and can be used as probes for a
CC variety of analyses such as gene expression analysis. Each probe
CC comprises 9 or more consecutive nucleotides from at least one of 14936
CC nucleotide sequences defined in the patent, or their perfect sense match,
CC sense mismatch, antisense match or antisense mismatch oligonucleotides.
CC The probes may be used in an array comprising at least 10 distinct
CC nucleic acid probes. The array is useful in monitoring gene expression
CC levels by hybridisation to a DNA library, in analysing genetic
CC variations, and in hybridising tag-labelled compounds. The probes are
CC useful for identifying family members of a gene. The probes are also
CC useful in in situ hybridisations, in screening cDNA or genomic libraries
CC (or derived subclones) for additional clones containing segments of DNA
CC that have been previously isolated and sequenced, in Southern, northern,
CC or dot-blot hybridisation of genomic DNA to identify or detect the
CC sequence of any gene or detect specific mutations in any gene, and in
CC mapping the 5' termini of mRNA molecules by primer extensions. The
CC nucleic acid sequences of the invention are also useful as PCR primers.
CC The invention provides a large collection of nucleic acid sequences
CC complementary to particular genes with a wide range of analytical uses.
CC ACH50865-ACH65260 represent the target sequences of the invention. Note:
CC The sequence data for this patent was obtained in electronic format
CC directly from the USPTO web site at seqdata.uspto.gov/psipsDIDEntry.html
XX
SQ Sequence 25 BP; 6 A; 4 C; 4 G; 11 T; 0 U; 0 Other;

Query Match 64.0%; Score 12.8; DB 8; Length 25;
Best Local Similarity 87.5%; Pred. No. 1.6e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 AACATCTATGTTTGGT 19
||| |||||||||| |
Db 2 AACTTCTATGTTTGCT 17

RESULT 20
AAS09822
ID AAS09822 standard; DNA; 30 BP.
XX
AC AAS09822;
XX
DT 24-OCT-2001 (first entry)
XX
DE Oat Beta-amyrin synthase sequencing primer 64.
XX
KW Oat; Beta-amyrin synthase; triterpenoid; palatability;
KW oxidosqualene cyclase; pathogen resistance; transgenic plant;
KW fungal disease; sequencing primer; ss.
XX
OS Avena strigosa.
XX
PN WO200146391-A2.
XX
PD 28-JUN-2001.
XX
PF 20-DEC-2000; 2000WO-GB004908.
XX
PR 22-DEC-1999; 99GB-00030394.
PR 16-AUG-2000; 2000GB-00020217.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Osbourn AE, Haralampidis K, Bryan GT;
XX
DR WPI; 2001-418055/44.
XX
PT Novel beta-amyrin synthase encoding nucleic acids useful for influencing
PT or affecting triterpene synthesis, and hence resistance to fungal
PT pathogen, taste, palatability or nutritional value of plants.
XX
PS Example 4; Page 60; 69pp; English.
XX
CC The sequence represents a primer used to sequence nucleic acids encoding
CC Oat Beta-amyrin synthase (an oxidosqualene cyclase). Beta-amyrin is a
CC triterpenoid responsible for paltablity to animals and resistance to
CC pathogens and predators. The beta-amyrin synthase encoding nucleic acid
CC is useful for producing a transgenic plant, by introducing a vector
CC containing it into a host cell, optionally causing or allowing
CC recombination between the vector and the host cell genome so as to
CC transform the host cell, and regenerating a plant from the transformed
CC plant cell. The DNA is also useful for identifying, cloning or
CC determining the presence of a nucleic acid in a sample and for
CC influencing or affecting the quantity or quality of triterpenoid
CC synthesis, preferably an oleanane-type triterpene saponin synthesis, in a
CC plant, such as altering resistance to a fungal pathogen e.g., an
CC ascomycete having a sterol-containing membrane, optionally selected from
CC Gaeumannomyces graminis vars tritici and avenae, Fusarium culmorus, F.
CC avanaceum, Stagonospora nodorum or S. avenae, taste, palatability and/or
CC nutritional value, of the plant, by causing or allowing expression of the
CC DNA within the cells of the plant, following an earlier step of
CC introducing the DNA into a cell or its ancestor. The DNA is also useful
CC for reducing the level of triterpenoids in the plant, by causing or
CC allowing transcription from an antisense molecule in the plant, allowing
CC transcription from the DNA, or its part such as to reduce beta-amyrin
CC synthase expression by co-suppression, use of a nucleic acid encoding a
CC ribozyme specific for the DNA
XX
SQ Sequence 30 BP; 7 A; 8 C; 4 G; 11 T; 0 U; 0 Other;

Query Match 64.0%; Score 12.8; DB 4; Length 30;
Best Local Similarity 87.5%; Pred. No. 1.6e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 ACATCTATGTTTGGTT 20
||||| ||||||| ||
Db 7 ACATCCATGTTTGTTT 22

RESULT 21
AAF83283
ID AAF83283 standard; DNA; 34 BP.
XX
AC AAF83283;
XX
DT 09-JUL-2001 (first entry)
XX
DE Human Chk1 DNA amplifying primer chk6w.
XX
KW Effector checkpoint protein kinase; Chk1; hyperproliferation; HIV;
KW cancer; cytostatic; anti HIV; gene therapy; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN EP1096014-A2.
XX
PD 02-MAY-2001.

```
XX
PF   31-OCT-2000; 2000EP-00123738.
XX
PR   01-NOV-1999;   99US-0162887P.
PR   14-DEC-1999;   99US-00460421.
XX
PA   (AGOU-) AGOURON PHARM INC.
XX
PI   Chen P,  Kan C,  Luo C,  Margosiak S,  O'connor P,  Tempczyk-Russel A;
PI   Nguyen B,  Sarup JC,  Gaur S,  Anderson MB,  Deng Y,  Lundgren K;
PI   Register J;
XX
DR   WPI; 2001-302195/32.
XX
PT   Novel isolated, soluble, catalytically active human effector checkpoint
PT   protein kinase, useful for screening inhibitors of hChk1 kinase, for
PT   treating hyperproliferative disorders such as HIV and cancer.
XX
PS   Example 2; Page 15; 169pp; English.
XX
CC   The invention relates to an isolated, soluble, catalytically active human
CC   effector checkpoint protein kinase (Chk1) polypeptide. Chk1 protein can
CC   be expressed by standard recombinant methodology. Chk1 is useful for
CC   screening for its inhibitors, used for treating hyperproliferative
CC   diseases, such as, HIV and cancer. The Chk1 DNA is useful for probes,
CC   primers, chemical intermediates, and in biological assays. Sequences
CC   AAF83283-290 represent PCR primers for amplifying the human Chk1 DNA
XX
SQ   Sequence 34 BP; 6 A; 7 C; 7 G; 14 T; 0 U; 0 Other;

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 34;
  Best Local Similarity   87.5%;  Pred. No. 1.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTT 16
              |||| ||||||| |||
Db          7 AGTACCATCTATCTTT 22


RESULT 22
ABZ27765/c
ID   ABZ27765 standard; DNA; 43 BP.
XX
AC   ABZ27765;
XX
DT   30-JAN-2003  (first entry)
XX
DE   Candida essential gene related knockout PCR primer SEQ ID NO 1712.
XX
KW   Fungus; yeast; tetracyclin; promoter; GRACE strain; biosynthesis;
KW   signal transduction; DNA replication; cell division; growth;
KW   proliferation; Candida albicans; fungicide; antifungal; PCR; primer; ss.
XX
OS   Candida albicans.
XX
PN   WO200253728-A2.
XX
PD   11-JUL-2002.
XX
PF   26-DEC-2001; 2001WO-US049486.
XX
PR   29-DEC-2000; 2000US-0259128P.
PR   20-FEB-2001; 2001US-00792024.
PR   22-AUG-2001; 2001US-0314050P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Roemer T,  Jiang B,  Boone C,  Bussey H,  Ohlsen KL;
XX
DR   WPI; 2002-566694/60.
XX
PT   Constructing strains for identifying gene products as effective targets
PT   for therapeutic intervention, by inactivating in the strain one allele of
PT   a gene and placing other allele of the gene under conditional expression.
XX
PS   Claim 76; SEQ ID NO 1712; 167pp + Sequence Listing; English.
XX
CC   The invention relates to constructing (M1) a strain of diploid fungal
CC   cells in which both alleles of a gene are modified, comprising modifying
CC   one allele by insertion or replacement by a cassette having an
CC   expressible selectable marker and modifying other allele by
CC   recombination, of a promoter replacement fragment with a heterologous
CC   promoter, so that expression of the second allele is regulated by the
CC   promoter. (M1) is useful for constructing a strain of diploid fungal
CC   cells in which both alleles of a gene are modified. The diploid fungal
CC   cells having both alleles modified are useful for identifying a gene that
CC   is essential to the survival or growth of a fungus, a gene that
CC   contributes to the virulence and/or pathogenicity of a fungus, a gene
CC   that contributes to the resistance of a diploid fungus to an antifungal
CC   agent, an antifungal agent that inhibits the growth of a diploid fungus
CC   and for identifying a therapeutic agent for treatment of a mammalian
CC   disease. (M1) is useful for identifying a compound which modulates the
CC   activity of a gene product, preferably enzymatic activity, carbon
CC   compound catabolism, biosynthetic, transporter, transcriptional,
CC   translational, signal transduction, DNA replication and cell division
CC   activity. The method is useful for identifying a compound having the
CC   ability to inhibit growth or proliferation of C. albicans cells and for
CC   treating infection by C. albicans. The present sequence is that of a PCR
CC   primer used in the method of the invention. Note: The sequence data for
CC   this patent is not represented in the printed specification but is based
CC   on sequence information supplied to Derwent by the European Patent Office
XX
SQ   Sequence 43 BP; 24 A; 4 C; 3 G; 12 T; 0 U; 0 Other;

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 43;
  Best Local Similarity   87.5%;  Pred. No. 1.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGGTT 20
              | ||||||||||| ||
Db         31 AGATCTATGTTTGTTT 16


RESULT 23
AAZ67540/c
ID   AAZ67540 standard; DNA; 47 BP.
XX
AC   AAZ67540;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human map-related biallelic marker SEQ ID NO:1887.
XX
KW   Human genome; biallelic marker; high density disequilibrium map;
KW   genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW   haplotyping; hybridisation; identification; characterisation; diagnosis;
KW   single nucleotide polymorphism; SNP; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(24,C)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO9954500-A2.
XX
PD   28-OCT-1999.
XX
PF   21-APR-1999;   99WO-IB000822.
XX
PR   21-APR-1998;   98US-0082614P.
PR   23-NOV-1998;   98US-0109732P.
XX
PA   (GEST ) GENSET.
XX
```

PI Cohen D, Blumenfeld M, Chumakov I;
XX
DR WPI; 2000-013267/01.
XX
PT Novel biallelic markers used to construct a high density disequilibrium
PT map of the human genome.
XX
PS Claim 1; Page 633; 2745pp; English.
XX
CC AAZ65654 to AAZ69578 represent human biallelic markers from the present
CC invention, which contain a polymorphic base at position 24 of their
CC nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
CC primers for the biallelic markers. The biallelic markers of the invention
CC have a variety of uses: they can be used for high density mapping of the
CC human genome, and in complex association studies and haplotyping studies
CC which are useful in determining the genetic basis for disease states.
CC Compositions and methods of the invention can also be useful for the
CC identification of the targets for the development of pharmaceutical
CC agents and diagnostic methods, as well as the characterisation of the
CC differential efficacious responses to and side effects from
CC pharmaceutical agents acting on a disease as well as other treatment.
CC N.B. The SEQ ID NOs 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297 and
CC 3367, are not actually given a sequence in the Sequence Listing from the
CC present invention
XX
SQ Sequence 47 BP; 22 A; 5 C; 4 G; 16 T; 0 U; 0 Other;

Query Match 64.0%; Score 12.8; DB 3; Length 47;
Best Local Similarity 87.5%; Pred. No. 1.7e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 ACATCTATGTTTGGTT 20
||||.|||||||.||
Db 43 ACATTTATGTTTGTTT 28

RESULT 24
ABL00484/c
ID ABL00484 standard; DNA; 51 BP.
XX
AC ABL00484;
XX
DT 05-MAR-2002 (first entry)
XX
DE Human silent noncoding SNP oligonucleotide SEQ ID NO:475.
XX
KW Human; single nucleotide polymorphism; SNP; polymorphism; cytostatic;
KW immunosuppressive; antiinflammatory; neuroprotective; antimicrobial;
KW autoimmune disease; inflammation; cancer; nervous system disease;
KW infection; polymorphic protein; ds.
XX
OS Homo sapiens.
XX
PN WO200138586-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US032311.
XX
PR 24-NOV-1999; 99US-0167383P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-355949/37.
XX
PT Isolated human nucleic acids comprising one or more single nucleotide
PT polymorphisms, useful for treating a subject suffering from a pathology,
PT e.g. autoimmune diseases, ascribed to the presence of a sequence
PT polymorphism.
XX
PS Claim 1; Page 391; 674pp; English.
XX
CC ABL00010 to ABL01104 represent human nucleic acid oligonucleotides
CC comprising one or more single nucleotide polymorphisms (SNPs). ABB56531
CC to ABB56903 represent human peptides encoded by some of the SNP
CC oligonucleotides. The sequences from the present invention can have
CC immunosuppressive, cytostatic, antiinflammatory, neuroprotective and
CC antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides
CC and antibodies from the present invention can be used for treating a
CC subject suffering from, at risk for, or suspected of, suffering from a
CC pathology ascribed to the presence of a sequence polymorphism. The
CC pathology may be autoimmune diseases, inflammation, cancer, diseases of
CC the nervous system, and infection by pathogenic microorganisms. The SNPs
CC are also useful for determining which forms of a characterised
CC polymorphism are present in individuals. The antibodies may be used in
CC the detection, quantitation and/or cellular or tissue localisation of a
CC polymorphic protein (e.g., for use in measuring levels of the polymorphic
CC protein within appropriate physiological samples)
XX
SQ Sequence 51 BP; 11 A; 12 C; 13 G; 15 T; 0 U; 0 Other;

Query Match 64.0%; Score 12.8; DB 5; Length 51;
Best Local Similarity 87.5%; Pred. No. 1.7e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 TAACATCTATGTTTGG 18
||||||||||.|||
Db 38 TAACATCTATGAGTGG 23

RESULT 25
AAZ44622/c
ID AAZ44622 standard; DNA; 58 BP.
XX
AC AAZ44622;
XX
DT 07-APR-2000 (first entry)
XX
DE Newcastle disease virus LaSota primer BGL5F2.
XX
KW Avian-paramyxovirus; infection; lentogenic; F protein; vaccine;
KW respiratory disease; gastrointestinal disease; poultry pathogen;
KW local immunity; primer; ss.
XX
OS Newcastle disease virus.
XX
PN WO9966045-A1.
XX
PD 23-DEC-1999.
XX
PF 17-JUN-1999; 99WO-NL000377.
XX
PR 19-JUN-1998; 98EP-00202054.
XX
PA (DIEN-) STICHTING DIENST LANDBOUWKUNDIG ONDERZOE.
XX
PI Peeters BPH, De Leeuw OS, Koch G, Gielkens ALJ;
XX
DR WPI; 2000-106102/09.
XX
PT New avian paramyxovirus cDNA, useful for production of vaccine against
PT Newcastle disease virus.
XX
PS Disclosure; Page 33; 115pp; English.
XX
CC This invention describes a novel avian-paramyxovirus cDNA (I) which
CC comprises a nucleic acid sequence corresponding to the 5' terminal end of
CC the genome of avian-paramyxovirus allowing the generation of an
CC infectious copy of avian-paramyxovirus. The cell line is useful for the
CC production of infectious lentogenic NDV (Newcastle Disease virus) without
CC the addition of exogenous proteolytic activity. Also it is possible to
CC generate a stable transfected cell line that expresses the wild-type F
CC protein in the virus envelope therefore providing infectious particles,
CC useful in the form of a vaccine, especially against respiratory and/or

```
CC   gastrointestinal diseases. NDV can be easily cultured to very high titers
CC   in embryonated eggs. Mass culture of embryonated eggs is relatively
CC   cheap. NDV vaccines are relatively stable and can be simply administered
CC   by mass application methods e.g. drinking water or by spraying or by
CC   aerosol formation. The natural route of infection is by the respiratory
CC   and/or gastrointestinal tract which are also the major routes of
CC   infection of many other poultry pathogens. NDV can induce local immunity
CC   despite the presence of circulating maternal antibody. AAZ44527-Z44609
CC   and AAZ44618-Z44650 represent primers used in the isolation of the NDV
CC   strain LaSota genome
XX
SQ   Sequence 58 BP; 21 A; 11 C; 11 G; 15 T; 0 U; 0 Other;

  Query Match             64.0%;  Score 12.8;  DB 3;  Length 58;
  Best Local Similarity   87.5%;  Pred. No. 1.7e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGGTT 20
              | ||||  |||||||||
Db         50 AAATCTTTGTTTGGTT 35


RESULT 26
AAF83273/c
ID   AAF83273 standard; DNA; 74 BP.
XX
AC   AAF83273;
XX
DT   09-JUL-2001  (first entry)
XX
DE   S. cerevisiae YOL077C knock-out mutant constructing primer UPTAG.
XX
KW   Germination; proliferation; essential gene; YDR141C; YDR091C; YOL026C;
KW   YOL034W; YOL077C; antifungal; fungal disease; YOL022C; antisense therapy;
KW   mutant; PCR primer; ss.
XX
OS   Saccharomyces cerevisiae.
XX
PN   US6221597-B1.
XX
PD   24-APR-2001.
XX
PF   21-MAY-1999;   99US-00315793.
XX
PR   21-MAY-1999;   99US-00315793.
XX
PA   (ROSE-) ROSETTA INPHARMATICS INC.
XX
PI   Roberts CJ;
XX
DR   WPI; 2001-315575/33.
XX
PT   Identifying antifungal compounds for treating fungal and proliferative
PT   diseases, by using yeast genes essential for germination and
PT   proliferation as targets.
XX
PS   Example 6; Fig 32; 91pp; English.
XX
CC   The invention relates to genes in S. cerevisiae which are essential for
CC   germination and proliferation. The essential genes (EG) such as YDR141C,
CC   YDR091C, YOL022C, YOL026C, YOL034W and YOL077C are used in a method for
CC   identifying potential antifungal compounds (Cp). The method comprises
CC   overexpressing the EG cells, isolating a subset of genes induced/
CC   repressed by overexpression of EG and determining effect of Cp on down/up
CC   -regulation of any subset of genes or contacting a protein encoded by EG
CC   with Cp and determining binding between them. Cp is identified as a
CC   potential antifungal Cp, if it downregulates a gene that is induced by
CC   overexpression of EG or if it upregulates the gene that is repressed by
CC   the overexpression of EG and if Cp binds to the protein encoded by EG.
CC   The method is useful for identifying novel antifungal compounds for
CC   treating fungal diseases and proliferative disorders in humans and non-
CC   human mammals, including monkeys and other primates, dogs, cats.
CC   Sequences AAF823273-280 represents PCR primers for the construction and
CC   analysis of S. cerevisiae YOL077C knock-out mutant
XX
SQ   Sequence 74 BP; 18 A; 20 C; 20 G; 16 T; 0 U; 0 Other;

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 74;
  Best Local Similarity   87.5%;  Pred. No. 1.7e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGGTT 20
              | ||||  |||||||||
Db         23 ACATCCATCTTTGGTT 8


RESULT 27
ACC73349
ID   ACC73349 standard; DNA; 20 BP.
XX
AC   ACC73349;
XX
DT   15-JUL-2003  (first entry)
XX
DE   M marinum - M ulcerans specific probe MAR-ULC-02.
XX
KW   Microarray; probe; Mycobacterium; antibiotic-resistance; genotyping; ss.
XX
OS   Mycobacterium marinum.
OS   Mycobacterium ulcerans.
XX
PN   WO2003031654-A1.
XX
PD   17-APR-2003.
XX
PF   09-OCT-2002; 2002WO-KR001885.
XX
PR   09-OCT-2001; 2001KR-00062125.
XX
PA   (SJHI-) SJ HIGHTECH CO LTD.
PA   (KIMC/) KIM C.
PA   (PARK/) PARK H.
XX
PI   Kim C,  Park H,  Jang H,  Song E;
XX
DR   WPI; 2003-403109/38.
XX
PT   Microarray for simultaneously genotyping Mycobacteria species,
PT   differentiating Mycobacterium tuberculosis strains and detecting
PT   antibiotic-resistant strains, comprises specific probes on a support.
XX
PS   Claim 12; Page 57; 76pp; English.
XX
CC   The invention relates to a microarray comprising a support, a first probe
CC   for genotyping Mycobacterium species, second probe for differentiating
CC   Mycobacterium tuberculosis strains, and a third probe for detecting
CC   antibiotic-resistant strains, where the probes are immobilized on the
CC   support. This sequence represents an example of the first probe used for
CC   genotyping Mycobacterium species. The array is useful for simultaneously
CC   genotyping Mycobacterium species, differentiating M. tuberculosis strains
CC   and detecting antibiotic-resistant strains
XX
SQ   Sequence 20 BP; 3 A; 5 C; 4 G; 8 T; 0 U; 0 Other;

  Query Match             63.0%;  Score 12.6;  DB 7;  Length 20;
  Best Local Similarity   78.9%;  Pred. No. 2e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGTT 20
              | ||||||| |||| | ||
Db          1 GCAACATCTCTGTTGGTTT 19


RESULT 28
AAX02618/c
ID   AAX02618 standard; DNA; 24 BP.
```

XX
AC AAX02618;
XX
DT 07-MAY-1999 (first entry)
XX
DE S. aureus SecA2 PCR primer #1.
XX
KW SecA2; ATPase subunit; preprotein translocase; diagnosis; infection;
KW treatment; inhibit; antagonist; agonist; disease; bacterial; cardiac;
KW respiratory tract; gastrointestinal; central nervous system; CNS; eye;
KW kidney; urinary tract; skin; bone; joint; bacterial adhesion; wound;
KW matrix proteins; body implant; PCR primer; ss.
XX
OS Synthetic.
OS Staphylococcus aureus.
XX
PN EP892064-A2.
XX
PD 20-JAN-1999.
XX
PF 01-JUN-1998; 98EP-00304311.
XX
PR 04-JUN-1997; 97US-00868699.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Warren R, O'dwyer K, Perry C;
XX
DR WPI; 1999-083581/08.
XX
PT New Staphylococcus aureus ATPase subunit of preprotein translocase
PT (SecA2) polypeptide and polynucleotide - useful as diagnostic reagents
PT and for prevention and treatment of Staphylococcus aureus infections,
PT including toxic shock syndrome and splenic abscess.
XX
PS Disclosure; Page 32; 33pp; English.
XX
CC This invention describes the isolation of a novel Staphylococcus aureus
CC ATPase subunit of preprotein translocase (SecA). SecA polypeptides and
CC polynucleotides are useful for diagnosing diseases related to over or
CC underexpression of SecA protein by identifying mutations in the SecA
CC gene, or determining SecA polypeptide or mRNA expression levels due to an
CC infection of an organism with the SecA gene. They can diagnose the stage
CC and type of infection. SecA polypeptides can be used in treatment to
CC inhibit (antagonist i.e. antibacterial drugs) or enhance (agonist) SecA
CC activity. Diseases diagnosed, prevented or treated include bacterial
CC infections, especially Staphylococcus aureus infections of the upper and
CC lower respiratory tract (e.g. otitis media, thyroiditis), cardiac (e.g.
CC infective endocarditis), gastrointestinal (e.g. secretory diarrhoea,
CC splenic abscess), CNS (e.g. cerebral abscess), eye (e.g. conjunctivitis,
CC keratitis), kidney and urinary tract (e.g. toxic shock syndrome), skin
CC (e.g. impetigo, wound infection), and bone and joint (e.g. septic
CC arthritis, osteomyelitis). SecA polypeptides, polynucleotides and their
CC (ant)agonists can prevent adhesion of bacteria to matrix proteins, and
CC are useful for use on wounds and body implants to prevent bacterial
CC infection
XX
SQ Sequence 24 BP; 11 A; 3 C; 4 G; 6 T; 0 U; 0 Other;

Query Match 63.0%; Score 12.6; DB 2; Length 24;
Best Local Similarity 78.9%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 GTAACATCTATGTTTGGTT 20
|||||||||| || |||
Db 23 GTAACATCTAGTTTATGTT 5

RESULT 29
ACK28029/c
ID ACK28029 standard; DNA; 25 BP.
XX
AC ACK28029;
XX
DT 14-OCT-2003 (first entry)
XX
DE Human microarray DNA oligonucleotide SEQ ID NO 128010.
XX
KW EST; ss; probe; expressed sequence tag; microarray; gene expression;
KW genetic variation; biallelic marker; polymorphism; human;
KW cross-species comparison.
XX
OS Homo sapiens.
XX
PN US2003104410-A1.
XX
PD 05-JUN-2003.
XX
PF 15-MAR-2002; 2002US-00098263.
XX
PR 16-MAR-2001; 2001US-0276759P.
XX
PA (AFFY-) AFFYMETRIX INC.
XX
PI Mittmann MP;
XX
DR WPI; 2003-567953/53.
XX
PT New array of nucleic acid probes, useful for in situ hybridization, in
PT Southern, Northern or dot-blot hybridization to identify or detect the
PT sequence or specific mutations of any gene.
XX
PS Claim 1; SEQ ID NO 128010; 9pp; English.
XX
CC The invention discloses a microarray comprising a plurality of nucleic
CC acid probes including one of 2,018,500 fully defined sequences, or its
CC perfect match, perfect mismatch, antisense match or antisense mismatch.
CC Also disclosed is a method of gene expression analysis. The array is used
CC in monitoring gene expression levels by hybridisation to a DNA library,
CC in analysis of genetic variation or in hybridisation of tag-labelled
CC compounds. The nucleic acid probes are specifically designed for analysis
CC of at least one target sequence. The method of analysis comprises
CC hybridising at least one or more nucleic acids to at least two or more
CC nucleic acid probes and detecting the hybridisation. The nucleic acid
CC probes are attached to a solid support. The analysis comprises monitoring
CC gene expression levels, identifying biallelic markers or polymorphisms,
CC or family members of a gene and a cross-species comparison. Each of the
CC nucleic acids further comprises a tag sequence. The array of nucleic acid
CC probes is useful in in situ hybridisation, in Southern, Northern or dot-
CC blot hybridisation to identify or detect the sequence or specific
CC mutations of any gene, in mapping the 5' termini of mRNA molecules by
CC primer extensions or in screening cDNA or genomic libraries or subclones
CC for additional subclones containing segments of DNA that have been
CC isolated and previously sequenced. The sequence presented is one of the
CC nucleic acid probes incorporated in the microarray. Note: The sequence
CC data for this patent can also be obtained in electronic format directly
CC from USPTO at seqdata.uspto.goc/sequence.html
XX
SQ Sequence 25 BP; 8 A; 8 C; 3 G; 6 T; 0 U; 0 Other;

Query Match 63.0%; Score 12.6; DB 8; Length 25;
Best Local Similarity 78.9%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 GTAACATCTATGTTTGGTT 20
|||| ||||||| ||| |
Db 24 GTAAAGTCTATGTATGGGT 6

RESULT 30
ACI29919/c
ID ACI29919 standard; DNA; 25 BP.
XX
AC ACI29919;
XX

DT 13-OCT-2003 (first entry)
XX
DE Human microarray DNA oligonucleotide SEQ ID NO 29910.
XX
KW EST; ss; probe; expressed sequence tag; microarray; gene expression;
KW genetic variation; biallelic marker; polymorphism; human;
KW cross-species comparison.
XX
OS Homo sapiens.
XX
PN US2003104410-A1.
XX
PD 05-JUN-2003.
XX
PF 15-MAR-2002; 2002US-00098263.
XX
PR 16-MAR-2001; 2001US-0276759P.
XX
PA (AFFY-) AFFYMETRIX INC.
XX
PI Mittmann MP;
XX
DR WPI; 2003-567953/53.
XX
PT New array of nucleic acid probes, useful for in situ hybridization, in
PT Southern, Northern or dot-blot hybridization to identify or detect the
PT sequence or specific mutations of any gene.
XX
PS Claim 1; SEQ ID NO 29910; 9pp; English.
XX
CC The invention discloses a microarray comprising a plurality of nucleic
CC acid probes including one of 2,018,500 fully defined sequences, or its
CC perfect match, perfect mismatch, antisense match or antisense mismatch.
CC Also disclosed is a method of gene expression analysis. The array is used
CC in monitoring gene expression levels by hybridisation to a DNA library,
CC in analysis of genetic variation or in hybridisation of tag-labelled
CC compounds. The nucleic acid probes are specifically designed for analysis
CC of at least one target sequence. The method of analysis comprises
CC hybridising at least one or more nucleic acids to at least two or more
CC nucleic acid probes and detecting the hybridisation. The nucleic acid
CC probes are attached to a solid support. The analysis comprises monitoring
CC gene expression levels, identifying biallelic markers or polymorphisms,
CC or family members of a gene and a cross-species comparison. Each of the
CC nucleic acids further comprises a tag sequence. The array of nucleic acid
CC probes is useful in in situ hybridisation, in Southern, Northern or dot-
CC blot hybridisation to identify or detect the sequence or specific
CC mutations of any gene, in mapping the 5' termini of mRNA molecules by
CC primer extensions or in screening cDNA or genomic libraries or subclones
CC for additional subclones containing segments of DNA that have been
CC isolated and previously sequenced. The sequence presented is one of the
CC nucleic acid probes incorporated in the microarray. Note: The sequence
CC data for this patent can also be obtained in electronic format directly
CC from USPTO at seqdata.uspto.goc/sequence.html
XX
SQ Sequence 25 BP; 7 A; 4 C; 5 G; 9 T; 0 U; 0 Other;

Query Match 63.0%; Score 12.6; DB 8; Length 25;
Best Local Similarity 78.9%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 AGTAACATCTATGTTTGGT 19
              ||||||||| | || || |
Db         19 AGTAACATCAAAGTCTGTT 1
```

RESULT 31
ACI36381/c
ID ACI36381 standard; DNA; 25 BP.
XX
AC ACI36381;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human microarray DNA oligonucleotide SEQ ID NO 36372.
XX
KW EST; ss; probe; expressed sequence tag; microarray; gene expression;
KW genetic variation; biallelic marker; polymorphism; human;
KW cross-species comparison.
XX
OS Homo sapiens.
XX
PN US2003104410-A1.
XX
PD 05-JUN-2003.
XX
PF 15-MAR-2002; 2002US-00098263.
XX
PR 16-MAR-2001; 2001US-0276759P.
XX
PA (AFFY-) AFFYMETRIX INC.
XX
PI Mittmann MP;
XX
DR WPI; 2003-567953/53.
XX
PT New array of nucleic acid probes, useful for in situ hybridization, in
PT Southern, Northern or dot-blot hybridization to identify or detect the
PT sequence or specific mutations of any gene.
XX
PS Claim 1; SEQ ID NO 36372; 9pp; English.
XX
CC The invention discloses a microarray comprising a plurality of nucleic
CC acid probes including one of 2,018,500 fully defined sequences, or its
CC perfect match, perfect mismatch, antisense match or antisense mismatch.
CC Also disclosed is a method of gene expression analysis. The array is used
CC in monitoring gene expression levels by hybridisation to a DNA library,
CC in analysis of genetic variation or in hybridisation of tag-labelled
CC compounds. The nucleic acid probes are specifically designed for analysis
CC of at least one target sequence. The method of analysis comprises
CC hybridising at least one or more nucleic acids to at least two or more
CC nucleic acid probes and detecting the hybridisation. The nucleic acid
CC probes are attached to a solid support. The analysis comprises monitoring
CC gene expression levels, identifying biallelic markers or polymorphisms,
CC or family members of a gene and a cross-species comparison. Each of the
CC nucleic acids further comprises a tag sequence. The array of nucleic acid
CC probes is useful in in situ hybridisation, in Southern, Northern or dot-
CC blot hybridisation to identify or detect the sequence or specific
CC mutations of any gene, in mapping the 5' termini of mRNA molecules by
CC primer extensions or in screening cDNA or genomic libraries or subclones
CC for additional subclones containing segments of DNA that have been
CC isolated and previously sequenced. The sequence presented is one of the
CC nucleic acid probes incorporated in the microarray. Note: The sequence
CC data for this patent can also be obtained in electronic format directly
CC from USPTO at seqdata.uspto.goc/sequence.html
XX
SQ Sequence 25 BP; 9 A; 4 C; 5 G; 7 T; 0 U; 0 Other;

Query Match 63.0%; Score 12.6; DB 8; Length 25;
Best Local Similarity 78.9%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 AGTAACATCTATGTTTGGT 19
              |||||| || |||| || |
Db         24 AGTAACTTCAATGTCTGTT 6
```

RESULT 32
ACI65204
ID ACI65204 standard; DNA; 25 BP.
XX
AC ACI65204;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human microarray DNA oligonucleotide SEQ ID NO 65195.
XX

KW EST; ss; probe; expressed sequence tag; microarray; gene expression;
KW genetic variation; biallelic marker; polymorphism; human;
KW cross-species comparison.
XX
OS Homo sapiens.
XX
PN US2003104410-A1.
XX
PD 05-JUN-2003.
XX
PF 15-MAR-2002; 2002US-00098263.
XX
PR 16-MAR-2001; 2001US-0276759P.
XX
PA (AFFY-) AFFYMETRIX INC.
XX
PI Mittmann MP;
XX
DR WPI; 2003-567953/53.
XX
PT New array of nucleic acid probes, useful for in situ hybridization, in
PT Southern, Northern or dot-blot hybridization to identify or detect the
PT sequence or specific mutations of any gene.
XX
PS Claim 1; SEQ ID NO 65195; 9pp; English.
XX
CC The invention discloses a microarray comprising a plurality of nucleic
CC acid probes including one of 2,018,500 fully defined sequences, or its
CC perfect match, perfect mismatch, antisense match or antisense mismatch.
CC Also disclosed is a method of gene expression analysis. The array is used
CC in monitoring gene expression levels by hybridisation to a DNA library,
CC in analysis of genetic variation or in hybridisation of tag-labelled
CC compounds. The nucleic acid probes are specifically designed for analysis
CC of at least one target sequence. The method of analysis comprises
CC hybridising at least one or more nucleic acids to at least two or more
CC nucleic acid probes and detecting the hybridisation. The nucleic acid
CC probes are attached to a solid support. The analysis comprises monitoring
CC gene expression levels, identifying biallelic markers or polymorphisms,
CC or family members of a gene and a cross-species comparison. Each of the
CC nucleic acids further comprises a tag sequence. The array of nucleic acid
CC probes is useful in in situ hybridisation, in Southern, Northern or dot-
CC blot hybridisation to identify or detect the sequence or specific
CC mutations of any gene, in mapping the 5' termini of mRNA molecules by
CC primer extensions or in screening cDNA or genomic libraries or subclones
CC for additional subclones containing segments of DNA that have been
CC isolated and previously sequenced. The sequence presented is one of the
CC nucleic acid probes incorporated in the microarray. Note: The sequence
CC data for this patent can also be obtained in electronic format directly
CC from USPTO at seqdata.uspto.goc/sequence.html
XX
SQ Sequence 25 BP; 4 A; 4 C; 8 G; 9 T; 0 U; 0 Other;

Query Match 63.0%; Score 12.6; DB 8; Length 25;
Best Local Similarity 78.9%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          2 GTAACATCTATGTTTGGTT 20
              |||||| || |||| |||
Db          2 GTAACAGGTAGGTTTCGTT 20
```

RESULT 33
ACK00121
ID ACK00121 standard; DNA; 25 BP.
XX
AC ACK00121;
XX
DT 14-OCT-2003 (first entry)
XX
DE Human microarray DNA oligonucleotide SEQ ID NO 100102.
XX
KW EST; ss; probe; expressed sequence tag; microarray; gene expression;
KW genetic variation; biallelic marker; polymorphism; human;
KW cross-species comparison.
XX
OS Homo sapiens.
XX
PN US2003104410-A1.
XX
PD 05-JUN-2003.
XX
PF 15-MAR-2002; 2002US-00098263.
XX
PR 16-MAR-2001; 2001US-0276759P.
XX
PA (AFFY-) AFFYMETRIX INC.
XX
PI Mittmann MP;
XX
DR WPI; 2003-567953/53.
XX
PT New array of nucleic acid probes, useful for in situ hybridization, in
PT Southern, Northern or dot-blot hybridization to identify or detect the
PT sequence or specific mutations of any gene.
XX
PS Claim 1; SEQ ID NO 100102; 9pp; English.
XX
CC The invention discloses a microarray comprising a plurality of nucleic
CC acid probes including one of 2,018,500 fully defined sequences, or its
CC perfect match, perfect mismatch, antisense match or antisense mismatch.
CC Also disclosed is a method of gene expression analysis. The array is used
CC in monitoring gene expression levels by hybridisation to a DNA library,
CC in analysis of genetic variation or in hybridisation of tag-labelled
CC compounds. The nucleic acid probes are specifically designed for analysis
CC of at least one target sequence. The method of analysis comprises
CC hybridising at least one or more nucleic acids to at least two or more
CC nucleic acid probes and detecting the hybridisation. The nucleic acid
CC probes are attached to a solid support. The analysis comprises monitoring
CC gene expression levels, identifying biallelic markers or polymorphisms,
CC or family members of a gene and a cross-species comparison. Each of the
CC nucleic acids further comprises a tag sequence. The array of nucleic acid
CC probes is useful in in situ hybridisation, in Southern, Northern or dot-
CC blot hybridisation to identify or detect the sequence or specific
CC mutations of any gene, in mapping the 5' termini of mRNA molecules by
CC primer extensions or in screening cDNA or genomic libraries or subclones
CC for additional subclones containing segments of DNA that have been
CC isolated and previously sequenced. The sequence presented is one of the
CC nucleic acid probes incorporated in the microarray. Note: The sequence
CC data for this patent can also be obtained in electronic format directly
CC from USPTO at seqdata.uspto.goc/sequence.html
XX
SQ Sequence 25 BP; 5 A; 7 C; 5 G; 8 T; 0 U; 0 Other;

Query Match 63.0%; Score 12.6; DB 8; Length 25;
Best Local Similarity 78.9%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 AGTAACATCTATGTTTGGT 19
              |||| |||||| |||  ||
Db          3 AGTACCATCTACGTTCCGT 21
```

RESULT 34
ACI27083
ID ACI27083 standard; DNA; 25 BP.
XX
AC ACI27083;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human microarray DNA oligonucleotide SEQ ID NO 27074.
XX
KW EST; ss; probe; expressed sequence tag; microarray; gene expression;
KW genetic variation; biallelic marker; polymorphism; human;
KW cross-species comparison.
XX

OS Homo sapiens.
XX
PN US2003104410-A1.
XX
PD 05-JUN-2003.
XX
PF 15-MAR-2002; 2002US-00098263.
XX
PR 16-MAR-2001; 2001US-0276759P.
XX
PA (AFFY-) AFFYMETRIX INC.
XX
PI Mittmann MP;
XX
DR WPI; 2003-567953/53.
XX
PT New array of nucleic acid probes, useful for in situ hybridization, in
PT Southern, Northern or dot-blot hybridization to identify or detect the
PT sequence or specific mutations of any gene.
XX
PS Claim 1; SEQ ID NO 27074; 9pp; English.
XX
CC The invention discloses a microarray comprising a plurality of nucleic
CC acid probes including one of 2,018,500 fully defined sequences, or its
CC perfect match, perfect mismatch, antisense match or antisense mismatch.
CC Also disclosed is a method of gene expression analysis. The array is used
CC in monitoring gene expression levels by hybridisation to a DNA library,
CC in analysis of genetic variation or in hybridisation of tag-labelled
CC compounds. The nucleic acid probes are specifically designed for analysis
CC of at least one target sequence. The method of analysis comprises
CC hybridising at least one or more nucleic acids to at least two or more
CC nucleic acid probes and detecting the hybridisation. The nucleic acid
CC probes are attached to a solid support. The analysis comprises monitoring
CC gene expression levels, identifying biallelic markers or polymorphisms,
CC or family members of a gene and a cross-species comparison. Each of the
CC nucleic acids further comprises a tag sequence. The array of nucleic acid
CC probes is useful in in situ hybridisation, in Southern, Northern or dot-
CC blot hybridisation to identify or detect the sequence or specific
CC mutations of any gene, in mapping the 5' termini of mRNA molecules by
CC primer extensions or in screening cDNA or genomic libraries or subclones
CC for additional subclones containing segments of DNA that have been
CC isolated and previously sequenced. The sequence presented is one of the
CC nucleic acid probes incorporated in the microarray. Note: The sequence
CC data for this patent can also be obtained in electronic format directly
CC from USPTO at seqdata.uspto.goc/sequence.html
XX
SQ Sequence 25 BP; 7 A; 3 C; 5 G; 10 T; 0 U; 0 Other;

Query Match 63.0%; Score 12.6; DB 8; Length 25;
Best Local Similarity 78.9%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 AGTAACATCTATGTTTGGT 19
              ||||| |  ||||||| ||
Db          1 AGTAATAGATATGTTTCGT 19
```

RESULT 35
ACI35745/c
ID ACI35745 standard; DNA; 25 BP.
XX
AC ACI35745;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human microarray DNA oligonucleotide SEQ ID NO 35736.
XX
KW EST; ss; probe; expressed sequence tag; microarray; gene expression;
KW genetic variation; biallelic marker; polymorphism; human;
KW cross-species comparison.
XX
OS Homo sapiens.
XX
PN US2003104410-A1.
XX
PD 05-JUN-2003.
XX
PF 15-MAR-2002; 2002US-00098263.
XX
PR 16-MAR-2001; 2001US-0276759P.
XX
PA (AFFY-) AFFYMETRIX INC.
XX
PI Mittmann MP;
XX
DR WPI; 2003-567953/53.
XX
PT New array of nucleic acid probes, useful for in situ hybridization, in
PT Southern, Northern or dot-blot hybridization to identify or detect the
PT sequence or specific mutations of any gene.
XX
PS Claim 1; SEQ ID NO 35736; 9pp; English.
XX
CC The invention discloses a microarray comprising a plurality of nucleic
CC acid probes including one of 2,018,500 fully defined sequences, or its
CC perfect match, perfect mismatch, antisense match or antisense mismatch.
CC Also disclosed is a method of gene expression analysis. The array is used
CC in monitoring gene expression levels by hybridisation to a DNA library,
CC in analysis of genetic variation or in hybridisation of tag-labelled
CC compounds. The nucleic acid probes are specifically designed for analysis
CC of at least one target sequence. The method of analysis comprises
CC hybridising at least one or more nucleic acids to at least two or more
CC nucleic acid probes and detecting the hybridisation. The nucleic acid
CC probes are attached to a solid support. The analysis comprises monitoring
CC gene expression levels, identifying biallelic markers or polymorphisms,
CC or family members of a gene and a cross-species comparison. Each of the
CC nucleic acids further comprises a tag sequence. The array of nucleic acid
CC probes is useful in in situ hybridisation, in Southern, Northern or dot-
CC blot hybridisation to identify or detect the sequence or specific
CC mutations of any gene, in mapping the 5' termini of mRNA molecules by
CC primer extensions or in screening cDNA or genomic libraries or subclones
CC for additional subclones containing segments of DNA that have been
CC isolated and previously sequenced. The sequence presented is one of the
CC nucleic acid probes incorporated in the microarray. Note: The sequence
CC data for this patent can also be obtained in electronic format directly
CC from USPTO at seqdata.uspto.goc/sequence.html
XX
SQ Sequence 25 BP; 9 A; 4 C; 5 G; 7 T; 0 U; 0 Other;

Query Match 63.0%; Score 12.6; DB 8; Length 25;
Best Local Similarity 78.9%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 AGTAACATCTATGTTTGGT 19
              |||||| || |||| || |
Db         24 AGTAACTTCAATGTCTGTT 6
```

RESULT 36
ACI25132/c
ID ACI25132 standard; DNA; 25 BP.
XX
AC ACI25132;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human microarray DNA oligonucleotide SEQ ID NO 25123.
XX
KW EST; ss; probe; expressed sequence tag; microarray; gene expression;
KW genetic variation; biallelic marker; polymorphism; human;
KW cross-species comparison.
XX
OS Homo sapiens.
XX
PN US2003104410-A1.
XX

PD 05-JUN-2003.
XX
PF 15-MAR-2002; 2002US-00098263.
XX
PR 16-MAR-2001; 2001US-0276759P.
XX
PA (AFFY-) AFFYMETRIX INC.
XX
PI Mittmann MP;
XX
DR WPI; 2003-567953/53.
XX
PT New array of nucleic acid probes, useful for in situ hybridization, in
PT Southern, Northern or dot-blot hybridization to identify or detect the
PT sequence or specific mutations of any gene.
XX
PS Claim 1; SEQ ID NO 25123; 9pp; English.
XX
CC The invention discloses a microarray comprising a plurality of nucleic
CC acid probes including one of 2,018,500 fully defined sequences, or its
CC perfect match, perfect mismatch, antisense match or antisense mismatch.
CC Also disclosed is a method of gene expression analysis. The array is used
CC in monitoring gene expression levels by hybridisation to a DNA library,
CC in analysis of genetic variation or in hybridisation of tag-labelled
CC compounds. The nucleic acid probes are specifically designed for analysis
CC of at least one target sequence. The method of analysis comprises
CC hybridising at least one or more nucleic acids to at least two or more
CC nucleic acid probes and detecting the hybridisation. The nucleic acid
CC probes are attached to a solid support. The analysis comprises monitoring
CC gene expression levels, identifying biallelic markers or polymorphisms,
CC or family members of a gene and a cross-species comparison. Each of the
CC nucleic acids further comprises a tag sequence. The array of nucleic acid
CC probes is useful in in situ hybridisation, in Southern, Northern or dot-
CC blot hybridisation to identify or detect the sequence or specific
CC mutations of any gene, in mapping the 5' termini of mRNA molecules by
CC primer extensions or in screening cDNA or genomic libraries or subclones
CC for additional subclones containing segments of DNA that have been
CC isolated and previously sequenced. The sequence presented is one of the
CC nucleic acid probes incorporated in the microarray. Note: The sequence
CC data for this patent can also be obtained in electronic format directly
CC from USPTO at seqdata.uspto.goc/sequence.html
XX
SQ Sequence 25 BP; 11 A; 6 C; 4 G; 4 T; 0 U; 0 Other;

Query Match 63.0%; Score 12.6; DB 8; Length 25;
Best Local Similarity 78.9%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          2 GTAACATCTATGTTTGGTT 20
              | |  |||||| |||||||
Db         23 GAAGGATCTATCTTTGGTT 5
```

RESULT 37
ACI35606
ID ACI35606 standard; DNA; 25 BP.
XX
AC ACI35606;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human microarray DNA oligonucleotide SEQ ID NO 35597.
XX
KW EST; ss; probe; expressed sequence tag; microarray; gene expression;
KW genetic variation; biallelic marker; polymorphism; human;
KW cross-species comparison.
XX
OS Homo sapiens.
XX
PN US2003104410-A1.
XX
PD 05-JUN-2003.
XX
PF 15-MAR-2002; 2002US-00098263.
XX
PR 16-MAR-2001; 2001US-0276759P.
XX
PA (AFFY-) AFFYMETRIX INC.
XX
PI Mittmann MP;
XX
DR WPI; 2003-567953/53.
XX
PT New array of nucleic acid probes, useful for in situ hybridization, in
PT Southern, Northern or dot-blot hybridization to identify or detect the
PT sequence or specific mutations of any gene.
XX
PS Claim 1; SEQ ID NO 35597; 9pp; English.
XX
CC The invention discloses a microarray comprising a plurality of nucleic
CC acid probes including one of 2,018,500 fully defined sequences, or its
CC perfect match, perfect mismatch, antisense match or antisense mismatch.
CC Also disclosed is a method of gene expression analysis. The array is used
CC in monitoring gene expression levels by hybridisation to a DNA library,
CC in analysis of genetic variation or in hybridisation of tag-labelled
CC compounds. The nucleic acid probes are specifically designed for analysis
CC of at least one target sequence. The method of analysis comprises
CC hybridising at least one or more nucleic acids to at least two or more
CC nucleic acid probes and detecting the hybridisation. The nucleic acid
CC probes are attached to a solid support. The analysis comprises monitoring
CC gene expression levels, identifying biallelic markers or polymorphisms,
CC or family members of a gene and a cross-species comparison. Each of the
CC nucleic acids further comprises a tag sequence. The array of nucleic acid
CC probes is useful in in situ hybridisation, in Southern, Northern or dot-
CC blot hybridisation to identify or detect the sequence or specific
CC mutations of any gene, in mapping the 5' termini of mRNA molecules by
CC primer extensions or in screening cDNA or genomic libraries or subclones
CC for additional subclones containing segments of DNA that have been
CC isolated and previously sequenced. The sequence presented is one of the
CC nucleic acid probes incorporated in the microarray. Note: The sequence
CC data for this patent can also be obtained in electronic format directly
CC from USPTO at seqdata.uspto.goc/sequence.html
XX
SQ Sequence 25 BP; 4 A; 5 C; 7 G; 9 T; 0 U; 0 Other;

Query Match 63.0%; Score 12.6; DB 8; Length 25;
Best Local Similarity 78.9%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          2 GTAACATCTATGTTTGGTT 20
              ||||||  || |||| |||
Db          1 GTAACAGGTAGGTTTCGTT 19
```

RESULT 38
AAZ18573/c
ID AAZ18573 standard; DNA; 27 BP.
XX
AC AAZ18573;
XX
DT 19-OCT-1999 (first entry)
XX
DE Primer for ASTH1 polmorphic microsatellite marker.
XX
KW ASTH1; asthma; human; chromosome 11p; ASTH1I; ASTHIJ; genetic locus; ss;
KW therapeutic; immunogen; polymorphism; PCR primer; microsatellite marker.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9937809-A1.
XX
PD 29-JUL-1999.
XX
PF 21-JAN-1998; 98WO-US001260.
XX

PR 21-JAN-1998; 98WO-US001260.
XX
PA (AXYS-) AXYS PHARM INC.
XX
PI Brooks-Wilson AR, Buckler A, Cardon L, Carey AH, Galvin M;
PI Miller A, North M;
XX
DR WPI; 1999-479058/40.
XX
PT Mammalian asthma related genes, useful for diagnosis of a predisposition
PT to development of asthma.
XX
PS Disclosure; Page 50; 195pp; English.
XX
CC The invention identifies a genetic locus ASTH1, associated with asthma,
CC mapped to human chromosome 11p. ASTH1I and ASTHIJ are genes present
CC within the locus, located close to each other on human chromosome 11p,
CC and have similar patterns of expression, and common sequence motifs. The
CC ASTH1 genes and fragments, encoded protein, genomic regulatory regions
CC and anti-ASTH1 antibodies are useful in the identification of individuals
CC predisposed to development of asthma, and for the modulation of gene
CC activity in vivo for prophylactic and therapeutic purposes. The ASTH1
CC protein is useful as an immunogen to raise specific antibodies, in drug
CC screening for compositions that mimic or modulate ASTH1 activity or
CC expression, including altered forms of ASTH1 protein, and as a
CC therapeutic. Sequences AAZ18510-Z18631 represent PCR primers for
CC polymorphic microsatellite markers in the ASTH1 region
XX
SQ Sequence 27 BP; 5 A; 6 C; 6 G; 10 T; 0 U; 0 Other;

Query Match 63.0%; Score 12.6; DB 2; Length 27;
Best Local Similarity 78.9%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AGTAACATCTATGTTTGGT 19
||||||||||| | ||||
Db 27 AGTAACATCTCAGCCTGGT 9

RESULT 39
AAA80480/c
ID AAA80480 standard; DNA; 27 BP.
XX
AC AAA80480;
XX
DT 22-NOV-2000 (first entry)
XX
DE ASTH1 polymorphic microsatellite marker AFMA154ZD1 primer, SEQ ID NO:223.
XX
KW ASTH1 locus; ASTH1I; ASTH1J; human; chromosome 11p; asthma;
KW bronchial hyperreactivity; ets family; transcription factor;
KW splice variant; genetic predisposition; polymorphism; antibody;
KW drug screening; prophylaxis; therapy; diagnosis;
KW polymorphic microsatellite marker flanking sequence;
KW batched analysis of genotypes; BAGs; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN US6087485-A.
XX
PD 11-JUL-2000.
XX
PF 21-JAN-1998; 98US-00009913.
XX
PR 21-JAN-1997; 97US-0035663P.
PR 01-JUL-1997; 97US-0051432P.
XX
PA (AXYS-) AXYS PHARM INC.
XX
PI Galvin M, Miller A, North M, Cardon L, Buckler A;
PI Brooks-Wilson AR, Carey AH;
XX
DR WPI; 2000-505109/45.
XX
PT New nucleic acids other than naturally occurring chromosomes encoding
PT ASTH1 protein, for e.g. screening compositions that modulate expression
PT or function of ASTH1 proteins or as diagnostics for genetic
PT predisposition to asthma.
XX
PS Example; Col 31-32; 131pp; English.
XX
CC The invention relates to the ASTH1 locus on the short arm of human
CC chromosome (11p). This locus comprises the ASTH1I and ASTH1J genes, which
CC are associated with a genetic predisposition to asthma and bronchial
CC hyperreactivity. The ASTH1I and ASTH1J genes are oriented in opposite
CC directions with the ASTH1 locus, and have similar patterns of expression
CC and common sequence motifs. They are both expressed in trachea, lung and
CC several other tissues. ASTH1I and ASTH1J are novel members of the ets
CC family of transcription factors, which have been implicated in the
CC activation of a variety of genes including the TCRa gene and cytokine
CC genes known to be important in the aetiology of asthma. Both ASTH1I and
CC ASTH1J mRNAs are alternatively spliced. Alternative splicing of
CC transcripts has no effect on the open reading frame of ASTH1J, as the
CC exons involved are all 5' to the start codon in exon b. In contrast,
CC alternative splicing of ASTH1I transcripts results in 3 different ASTH1I
CC isoforms. The invention also encompasses mouse asth1j protein. The ASTH1
CC nucleic acids are useful as diagnostics to identify a hereditary
CC predisposition to asthma, as probes for identifying ASTH1 related genes,
CC for identifying expression of the gene in a biological specimen, and for
CC generating genetically modified non-human animals or site specific gene
CC modifications in cell lines. The encoded ASTH1 proteins are useful as
CC immunogens to raise specific antibodies; in drug screening for
CC compositions that mimic or modulate activity or expression of ASTH1I
CC and/or ASTH1J (including altered forms of these proteins); and as a
CC therapeutic. The ASTH1 genes or fragments thereof, encoded proteins,
CC ASTH1 genomic regulatory regions, and anti-ASTH1I and anti-ASTH1J
CC antibodies are useful in the identification of individuals predisposed to
CC development of asthma, and for modulation of gene activity in vivo for
CC prophylactic and therapeutic purposes. The intact ASTH1I or ASTH1J
CC proteins or active fragments thereof may be used to modulate or reduce
CC bronchial hyperreactivity. Sequences AAA80417-A80538 represent sequences
CC flanking polymorphic microsatellite markers in the ASTH1 region, which
CC were also used as PCR primers for amplification of the markers for
CC batched analysis of genotypes (BAGs)
XX
SQ Sequence 27 BP; 5 A; 6 C; 6 G; 10 T; 0 U; 0 Other;

Query Match 63.0%; Score 12.6; DB 3; Length 27;
Best Local Similarity 78.9%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AGTAACATCTATGTTTGGT 19
||||||||||| | ||||
Db 27 AGTAACATCTCAGCCTGGT 9

RESULT 40
ACC72239/c
ID ACC72239 standard; DNA; 27 BP.
XX
AC ACC72239;
XX
DT 07-JUL-2003 (first entry)
XX
DE Forward Ag7094 PCR primer.
XX
KW Human; NOV; antidiabetic; anorectic; antibacterial; virucide;
KW immunomodulator; cytostatic; nootropic; neuroprotective;
KW antiparkinsonian; antilipaemic; gene therapy; metabolic disorder;
KW diabetes; obesity; infection; cachexia; cancer; PCR; primer;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorder; dyslipidaemia; ss.
XX
OS Homo sapiens.
XX
PN WO2003029423-A2.

XX
PD 10-APR-2003.
XX
PF 02-OCT-2002; 2002WO-US031358.
XX
PR 02-OCT-2001; 2001US-0326483P.
PR 05-OCT-2001; 2001US-0327342P.
PR 09-OCT-2001; 2001US-0327917P.
PR 09-OCT-2001; 2001US-0328029P.
PR 09-OCT-2001; 2001US-0328044P.
PR 09-OCT-2001; 2001US-0328056P.
PR 12-OCT-2001; 2001US-0328849P.
PR 15-OCT-2001; 2001US-0329414P.
PR 17-OCT-2001; 2001US-0330142P.
PR 22-OCT-2001; 2001US-0341058P.
PR 24-OCT-2001; 2001US-0339266P.
PR 24-OCT-2001; 2001US-0343629P.
PR 29-OCT-2001; 2001US-0349575P.
PR 01-NOV-2001; 2001US-0346357P.
PR 12-APR-2002; 2002US-0371972P.
PR 12-APR-2002; 2002US-0371980P.
PR 17-APR-2002; 2002US-0373261P.
PR 19-APR-2002; 2002US-0373805P.
PR 23-APR-2002; 2002US-0374738P.
PR 16-MAY-2002; 2002US-0381101P.
PR 17-MAY-2002; 2002US-0381635P.
PR 29-MAY-2002; 2002US-0383830P.
PR 01-OCT-2002; 2002US-00262839.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E;
PI Edinger SR, Ellerman K, Gerlach VL, Gorman L, Guo X, Ji W;
PI Kekuda R, Leach MD, Li L, Miller CE, Patturajan M, Rieger DK;
PI Rothenberg ME, Shimkets RA, Smithson G, Spytek KA, Taupier RJ;
PI Vernet CAM, Voss EZ, Zerhusen BD, Zhong M;
XX
DR WPI; 2003-381625/36.
XX
PT NOVX polypeptides and nucleic acids useful for diagnosing, preventing or
PT treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or
PT dyslipidemia, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Example C; Page 360; 487pp; English.
XX
CC The present invention relates to novel human NOV proteins and their
CC coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV
CC proteins are useful in manufacturing a medicament for treating a syndrome
CC associated with a human disease. The NOV proteins and coding sequences
CC may be used to diagnose, treat or prevent metabolic disorders such as
CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative
CC disorders such as Alzheimer's disease or Parkinson's disease, immune
CC disorders, haematopoietic disorders and various dyslipidaemias. The
CC present sequence is a PCR primer, used in an example from the invention
XX
SQ Sequence 27 BP; 10 A; 5 C; 4 G; 8 T; 0 U; 0 Other;

Query Match 63.0%; Score 12.6; DB 7; Length 27;
Best Local Similarity 78.9%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AGTAACATCTATGTTTGGT 19
|| ||||||| ||| ||||
Db 23 ATTAACATCATTGTATGGT 5

RESULT 41
AAQ37473
ID AAQ37473 standard; DNA; 34 BP.
XX
AC AAQ37473;
XX
DT 25-MAR-2003 (revised)
DT 03-JUL-1993 (first entry)
XX
DE Sequence of PCR primer which corresp. to the sense strand for AAs 76-87
DE of ciliary neurotrophic factor (CNTF).
XX
KW Neurotrophic factor; neuro-degenerative disease; therapy; ss.
XX
OS Synthetic.
XX
PN WO9303758-A1.
XX
PD 04-MAR-1993.
XX
PF 21-AUG-1992; 92WO-US007070.
XX
PR 23-AUG-1991; 91US-00749446.
XX
PA (SCIO-) SCIOS NOVA INC.
XX
PI Higaki JN, Tischer EG, Cordell B, Thompson SA;
XX
DR WPI; 1993-093724/11.
XX
PT Homogeneous neurotrophic factor for treating neuro-degenerative diseases,
PT e.g. Parkinsonism - comprise specified polypeptide sequence lacking micro
PT -heterogeneity associated with a related native sequence factor.
XX
PS Example; Page 14; 34pp; English.
XX
CC cDNA sequence AAQ37477,encoding a neutrotrophic factor(NF),was derived
CC from the native ciliary neurotrophic factor (CNTF) (e.g. EP-385060) by
CC PCR amplification. In order to obtain a DNA sequence encoding the NF(1-
CC 180), a portion of the DNA sequence encoding AAs 76-180 of full-length
CC CNTF was amplified using PCR from a bacterial expression vector, plasmid
CC pSP18 and synthetic PCR primers AAQ37473 and AAQ37474. (Updated on 25-MAR
CC -2003 to correct PN field.)
XX
SQ Sequence 34 BP; 5 A; 9 C; 5 G; 15 T; 0 U; 0 Other;

Query Match 63.0%; Score 12.6; DB 2; Length 34;
Best Local Similarity 78.9%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 GTAACATCTATGTTTGGTT 20
||| | || |||||| |||
Db 15 GTACCTTCCATGTTTTGTT 33

RESULT 42
ABK91114/c
ID ABK91114 standard; DNA; 50 BP.
XX
AC ABK91114;
XX
DT 05-NOV-2002 (first entry)
XX
DE 50 bp spacer DNA used in aptamer.
XX
KW Aptamer; ss; sequencing.
XX
OS Synthetic.
XX
PN WO200244195-A2.
XX
PD 06-JUN-2002.
XX
PF 28-NOV-2001; 2001WO-JP010400.
XX
PR 28-NOV-2000; 2000US-0253097P.
XX
PA (RIKE) RIKEN KK.
PA (HAYA/) HAYASHIZAKI Y.

XX
PI Hayashizaki Y;
XX
DR WPI; 2002-608230/65.
XX
PT New aptamer comprising one base capable of base pairing and different
PT from the standard Watson-Crick base, useful for isolating a specific
PT ligand from a pool of ligands.
XX
PS Example 4; Page 23; 56pp; English.
XX
CC This invention relates to novel isolated aptamers comprising at least one
CC base capable of base pairing and different from the standard Watson-Crick
CC (W-C) bases. The invention also comprises a method for sequencing nucleic
CC acids. The aptomers of the invention are useful for isolating a specific
CC ligand from a pool of ligands, by providing at least one specific
CC aptamer, mixing it with a pool of ligands, and recovering the specific
CC ligand bound to specific aptamer. The aptomers of the invention are
CC useful for detection of specific ligand from a biological sample, by
CC selecting at least one specific aptamer, capable of binding to a specific
CC ligand from a biological sample, mixing the at least one specific aptamer
CC with a biological sample to allow binding of the ligand to the at least
CC one aptamer, and detecting the presence and/or quantity of the specific
CC ligand from the biological sample bound to at least one aptamer. The
CC aptamer of the invention is useful as a drug and for therapeutic
CC treatment. The present sequence represents a 50 bp spacer oligonucleotide
CC used in the construction of an aptamer of the invention
XX
SQ Sequence 50 BP; 15 A; 14 C; 11 G; 10 T; 0 U; 0 Other;

Query Match 63.0%; Score 12.6; DB 6; Length 50;
Best Local Similarity 78.9%; Pred. No. 2.1e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 AGTAACATCTATGTTTGGT 19
              |||  |||||||||  |||
Db         21 AGTGTCATCTATGTCGGGT 3
```

RESULT 43
ABZ00880
ID ABZ00880 standard; DNA; 50 BP.
XX
AC ABZ00880;
XX
DT 09-JAN-2003 (first entry)
XX
DE Human leukocyte gene expression profiling probe SEQ ID NO 871.
XX
KW T7; leukocyte; gene expression profiling; allograft rejection;
KW atherosclerosis; congestive heart failure; systemic lupus erythematosus;
KW rheumatoid arthritis; osteoarthritis; cytomegalovirus; infection; probe;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200257414-A2.
XX
PD 25-JUL-2002.
XX
PF 22-OCT-2001; 2001WO-US047856.
XX
PR 20-OCT-2000; 2000US-0241994P.
PR 08-JUN-2001; 2001US-0296764P.
XX
PA (BIOC-) BIOCARDIA INC.
XX
PI Wohlgemuth J, Fry K, Matcuk G, Altman P, Prentice J, Phillips J;
PI Ly N, Woodward R, Quertermous T, Johnson F;
XX
DR WPI; 2002-636525/68.
XX
PT New system for leukocyte expression profiling, diagnosing a disease, or
PT monitoring (the rate of) progression of a disease, e.g. atherosclerosis
PT or congestive heart failure, comprises diagnostic oligonucleotides.
XX
PS Claim 1; Page 352; 0pp; English.
XX
CC The invention relates to a system for detecting gene expression, which
CC comprises one or two isolated DNA molecules that detect expression of a
CC gene, where the gene corresponds to any of 8143 oligonucleotides
CC (ABZ00010-ABZ08152) each having 50 base pairs (bp). The system is useful
CC for leukocyte expression profiling. It is particularly useful for
CC diagnosing a disease, monitoring (rate of) progression of a disease,
CC predicting therapeutic outcome, determining prognosis for a patient,
CC predicting disease complications in an individual or monitoring response
CC to treatment in an individual. The diseases include cardiac allograft
CC rejection, kidney allograft rejection, liver allograft rejection,
CC atherosclerosis, congestive heart failure, systemic lupus erythematosus,
CC rheumatoid arthritis, osteoarthritis or cytomegalovirus infection
XX
SQ Sequence 50 BP; 12 A; 10 C; 11 G; 17 T; 0 U; 0 Other;

Query Match 63.0%; Score 12.6; DB 6; Length 50;
Best Local Similarity 78.9%; Pred. No. 2.1e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 AGTAACATCTATGTTTGGT 19
              || |||| |  |||||||||
Db         28 AGAAACACCCTTGTTTGGT 46
```

RESULT 44
ABZ03294/c
ID ABZ03294 standard; DNA; 50 BP.
XX
AC ABZ03294;
XX
DT 09-JAN-2003 (first entry)
XX
DE Human leukocyte gene expression profiling probe SEQ ID NO 3285.
XX
KW T7; leukocyte; gene expression profiling; allograft rejection;
KW atherosclerosis; congestive heart failure; systemic lupus erythematosus;
KW rheumatoid arthritis; osteoarthritis; cytomegalovirus; infection; probe;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200257414-A2.
XX
PD 25-JUL-2002.
XX
PF 22-OCT-2001; 2001WO-US047856.
XX
PR 20-OCT-2000; 2000US-0241994P.
PR 08-JUN-2001; 2001US-0296764P.
XX
PA (BIOC-) BIOCARDIA INC.
XX
PI Wohlgemuth J, Fry K, Matcuk G, Altman P, Prentice J, Phillips J;
PI Ly N, Woodward R, Quertermous T, Johnson F;
XX
DR WPI; 2002-636525/68.
XX
PT New system for leukocyte expression profiling, diagnosing a disease, or
PT monitoring (the rate of) progression of a disease, e.g. atherosclerosis
PT or congestive heart failure, comprises diagnostic oligonucleotides.
XX
PS Claim 1; Page 432; 0pp; English.
XX
CC The invention relates to a system for detecting gene expression, which
CC comprises one or two isolated DNA molecules that detect expression of a
CC gene, where the gene corresponds to any of 8143 oligonucleotides
CC (ABZ00010-ABZ08152) each having 50 base pairs (bp). The system is useful
CC for leukocyte expression profiling. It is particularly useful for

CC diagnosing a disease, monitoring (rate of) progression of a disease,
CC predicting therapeutic outcome, determining prognosis for a patient,
CC predicting disease complications in an individual or monitoring response
CC to treatment in an individual. The diseases include cardiac allograft
CC rejection, kidney allograft rejection, liver allograft rejection,
CC atherosclerosis, congestive heart failure, systemic lupus erythematosus,
CC rheumatoid arthritis, osteoarthritis or cytomegalovirus infection
XX
SQ Sequence 50 BP; 13 A; 11 C; 11 G; 15 T; 0 U; 0 Other;

Query Match 63.0%; Score 12.6; DB 6; Length 50;
Best Local Similarity 78.9%; Pred. No. 2.1e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AGTAACATCTATGTTTGGT 19
||||||| ||||| ||
Db 34 AGTAACATGAATGTTGTGT 16

RESULT 45
ABZ28986
ID ABZ28986 standard; DNA; 55 BP.
XX
AC ABZ28986;
XX
DT 30-JAN-2003 (first entry)
XX
DE Candida gene related tetracyclin promoter PCR primer SEQ ID NO 3069.
XX
KW Fungus; yeast; tetracyclin; promoter; GRACE strain; biosynthesis;
KW signal transduction; DNA replication; cell division; growth;
KW proliferation; Candida albicans; fungicide; antifungal; PCR; primer; ss.
XX
OS Candida albicans.
XX
PN WO200253728-A2.
XX
PD 11-JUL-2002.
XX
PF 26-DEC-2001; 2001WO-US049486.
XX
PR 29-DEC-2000; 2000US-0259128P.
PR 20-FEB-2001; 2001US-00792024.
PR 22-AUG-2001; 2001US-0314050P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;
XX
DR WPI; 2002-566694/60.
XX
PT Constructing strains for identifying gene products as effective targets
PT for therapeutic intervention, by inactivating in the strain one allele of
PT a gene and placing other allele of the gene under conditional expression.
XX
PS Claim 76; SEQ ID NO 3069; 167pp + Sequence Listing; English.
XX
CC The invention relates to constructing (M1) a strain of diploid fungal
CC cells in which both alleles of a gene are modified, comprising modifying
CC one allele by insertion or replacement by a cassette having an
CC expressible selectable marker and modifying other allele by
CC recombination, of a promoter replacement fragment with a heterologous
CC promoter, so that expression of the second allele is regulated by the
CC promoter. (M1) is useful for constructing a strain of diploid fungal
CC cells in which both alleles of a gene are modified. The diploid fungal
CC cells having both alleles modified are useful for identifying a gene that
CC is essential to the survival or growth of a fungus, a gene that
CC contributes to the virulence and/or pathogenicity of a fungus, a gene
CC that contributes to the resistance of a diploid fungus to an antifungal
CC agent, an antifungal agent that inhibits the growth of a diploid fungus
CC and for identifying a therapeutic agent for treatment of a mammalian
CC disease. (M1) is useful for identifying a compound which modulates the
CC activity of a gene product, preferably enzymatic activity, carbon
CC compound catabolism, biosynthetic, transporter, transcriptional,
CC translational, signal transduction, DNA replication and cell division
CC activity. The method is useful for identifying a compound having the
CC ability to inhibit growth or proliferation of C. albicans cells and for
CC treating infection by C. albicans. The present sequence is that of a PCR
CC primer used in the method of the invention. Note: The sequence data for
CC this patent is not represented in the printed specification but is based
CC on sequence information supplied to Derwent by the European Patent Office
XX
SQ Sequence 55 BP; 11 A; 8 C; 8 G; 28 T; 0 U; 0 Other;

Query Match 63.0%; Score 12.6; DB 6; Length 55;
Best Local Similarity 78.9%; Pred. No. 2.1e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 GTAACATCTATGTTTGGTT 20
| ||| ||| |||||| ||
Db 22 GAAACTTCTTTGTTTGTTT 40

Search completed: September 23, 2004, 15:59:49
Job time : 190 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 23, 2004, 15:56:54 ; Search time 220 Seconds
(without alignments)
460.450 Million cell updates/sec

Title: US-10-798-923A-36
Perfect score: 20
Sequence: 1 agtaacatctatgtttggtt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3337386 seqs, 2532474682 residues

Total number of hits satisfying chosen parameters: 2022542

Minimum DB seq length: 0
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	15.2	76.0	37	15	US-10-059-273-6	Sequence 6, Appli
	2	14.4	72.0	25	15	US-10-098-263B-60525	Sequence 60525, A
	3	13.8	69.0	33	13	US-09-842-776A-46	Sequence 46, Appl
c	4	13.8	69.0	77	16	US-10-428-339-5	Sequence 5, Appli
c	5	13.6	68.0	20	17	US-10-688-706-2763	Sequence 2763, Ap
c	6	13.6	68.0	28	15	US-10-043-639A-3	Sequence 3, Appli
c	7	13.6	68.0	73	10	US-09-911-132A-27	Sequence 27, Appl
c	8	13.4	67.0	25	15	US-10-098-263B-17106	Sequence 17106, A
c	9	13.2	66.0	20	9	US-09-969-373-3917	Sequence 3917, Ap
c	10	13.2	66.0	20	17	US-10-688-706-2863	Sequence 2863, Ap
c	11	13.2	66.0	20	17	US-10-688-706-2868	Sequence 2868, Ap
	12	13.2	66.0	24	9	US-09-753-143-14	Sequence 14, Appl
	13	13.2	66.0	65	15	US-10-032-585-3574	Sequence 3574, Ap
	14	12.8	64.0	17	10	US-09-877-478-145	Sequence 145, App
	15	12.8	64.0	17	10	US-09-877-478-848	Sequence 848, App
	16	12.8	64.0	17	13	US-10-342-902-145	Sequence 145, App
	17	12.8	64.0	17	13	US-10-342-902-848	Sequence 848, App
	18	12.8	64.0	17	17	US-10-669-841-145	Sequence 145, App
	19	12.8	64.0	17	17	US-10-669-841-848	Sequence 848, App
c	20	12.8	64.0	20	17	US-10-688-706-3025	Sequence 3025, Ap
c	21	12.8	64.0	20	17	US-10-688-706-3034	Sequence 3034, Ap
	22	12.8	64.0	25	15	US-10-215-112-6520	Sequence 6520, Ap
	23	12.8	64.0	25	15	US-10-215-112-10981	Sequence 10981, A
	24	12.8	64.0	25	15	US-10-098-263B-1523	Sequence 1523, Ap
	25	12.8	64.0	25	15	US-10-098-263B-60526	Sequence 60526, A
c	26	12.8	64.0	25	17	US-10-775-169-3303	Sequence 3303, Ap
c	27	12.8	64.0	25	17	US-10-775-169-3304	Sequence 3304, Ap
	28	12.8	64.0	27	15	US-10-002-623-548	Sequence 548, App
	29	12.8	64.0	30	15	US-10-168-445-216	Sequence 216, App
	30	12.8	64.0	34	16	US-10-353-274-3	Sequence 3, Appli
c	31	12.8	64.0	36	13	US-10-027-632-58587	Sequence 58587, A
c	32	12.8	64.0	36	13	US-10-027-632-58595	Sequence 58595, A
c	33	12.8	64.0	36	16	US-10-027-632-58587	Sequence 58587, A
c	34	12.8	64.0	36	16	US-10-027-632-58595	Sequence 58595, A
c	35	12.8	64.0	43	15	US-10-032-585-1712	Sequence 1712, Ap
c	36	12.8	64.0	47	16	US-10-349-143-1887	Sequence 1887, Ap
c	37	12.6	63.0	20	17	US-10-688-706-2716	Sequence 2716, Ap
c	38	12.6	63.0	25	15	US-10-098-263B-25123	Sequence 25123, A
	39	12.6	63.0	25	15	US-10-098-263B-27074	Sequence 27074, A
c	40	12.6	63.0	25	15	US-10-098-263B-29910	Sequence 29910, A
	41	12.6	63.0	25	15	US-10-098-263B-35597	Sequence 35597, A
c	42	12.6	63.0	25	15	US-10-098-263B-35736	Sequence 35736, A
c	43	12.6	63.0	25	15	US-10-098-263B-36372	Sequence 36372, A
	44	12.6	63.0	25	15	US-10-098-263B-65195	Sequence 65195, A
	45	12.6	63.0	25	15	US-10-098-263B-100102	Sequence 100102,
c	46	12.6	63.0	25	15	US-10-098-263B-128010	Sequence 128010,
c	47	12.6	63.0	27	13	US-10-262-839-272	Sequence 272, App
	48	12.6	63.0	31	16	US-10-379-981-7	Sequence 7, Appli
	49	12.6	63.0	50	16	US-10-131-827-871	Sequence 871, App
c	50	12.6	63.0	50	16	US-10-131-827-3285	Sequence 3285, Ap
c	51	12.6	63.0	50	17	US-10-432-991-15	Sequence 15, Appl
	52	12.6	63.0	55	15	US-10-032-585-3069	Sequence 3069, Ap
c	53	12.6	63.0	60	10	US-09-908-975-18762	Sequence 18762, A
	54	12.6	63.0	65	10	US-09-908-975-3536	Sequence 3536, Ap
	55	12.6	63.0	65	10	US-09-908-975-23739	Sequence 23739, A
c	56	12.4	62.0	20	17	US-10-688-706-2717	Sequence 2717, Ap
c	57	12.4	62.0	20	17	US-10-688-706-2787	Sequence 2787, Ap
	58	12.4	62.0	25	15	US-10-215-112-1882	Sequence 1882, Ap
	59	12.4	62.0	25	15	US-10-215-112-6264	Sequence 6264, Ap
	60	12.4	62.0	25	15	US-10-215-112-10345	Sequence 10345, A
	61	12.4	62.0	25	15	US-10-098-263B-61987	Sequence 61987, A
	62	12.4	62.0	25	15	US-10-098-263B-98886	Sequence 98886, A
	63	12.4	62.0	25	15	US-10-098-263B-99511	Sequence 99511, A
c	64	12.4	62.0	43	13	US-10-027-632-175446	Sequence 175446,
c	65	12.4	62.0	43	16	US-10-027-632-175446	Sequence 175446,
	66	12.4	62.0	50	16	US-10-131-827-5074	Sequence 5074, Ap
	67	12.4	62.0	60	10	US-09-908-975-12895	Sequence 12895, A
c	68	12.4	62.0	60	10	US-09-908-975-13693	Sequence 13693, A
	69	12.4	62.0	65	10	US-09-908-975-4306	Sequence 4306, Ap
	70	12.2	61.0	21	9	US-09-810-993-26	Sequence 26, Appl
	71	12.2	61.0	21	15	US-10-251-210-26	Sequence 26, Appl
	72	12.2	61.0	23	15	US-10-132-080-21	Sequence 21, Appl
	73	12.2	61.0	23	15	US-10-125-690-19	Sequence 19, Appl
	74	12.2	61.0	23	15	US-10-100-556-21	Sequence 21, Appl
	75	12.2	61.0	23	15	US-10-100-218-21	Sequence 21, Appl
	76	12.2	61.0	23	15	US-10-134-296-21	Sequence 21, Appl
	77	12.2	61.0	23	15	US-10-141-533-21	Sequence 21, Appl
	78	12.2	61.0	23	15	US-10-072-611-19	Sequence 19, Appl
	79	12.2	61.0	23	15	US-10-135-185-21	Sequence 21, Appl
	80	12.2	61.0	23	15	US-10-125-693-19	Sequence 19, Appl
	81	12.2	61.0	23	15	US-10-164-854-21	Sequence 21, Appl
	82	12.2	61.0	23	15	US-10-087-996-21	Sequence 21, Appl
	83	12.2	61.0	23	15	US-10-100-230-21	Sequence 21, Appl
	84	12.2	61.0	23	15	US-10-163-598-2	Sequence 2, Appli
	85	12.2	61.0	23	15	US-10-100-272-21	Sequence 21, Appl
	86	12.2	61.0	23	15	US-10-325-466-21	Sequence 21, Appl
	87	12.2	61.0	23	15	US-10-151-467A-21	Sequence 21, Appl

88	12.2	61.0	23	15	US-10-124-747-21	Sequence 21, Appl
89	12.2	61.0	23	15	US-10-307-204-21	Sequence 21, Appl
90	12.2	61.0	23	17	US-10-607-848-21	Sequence 21, Appl
91	12.2	61.0	23	17	US-10-307-210-21	Sequence 21, Appl
92	12.2	61.0	25	15	US-10-098-263B-19514	Sequence 19514, A
93	12.2	61.0	25	15	US-10-098-263B-108355	Sequence 108355,
94	12.2	61.0	25	15	US-10-098-263B-113206	Sequence 113206,
95	12.2	61.0	25	17	US-10-775-169-2171	Sequence 2171, Ap
96	12.2	61.0	25	17	US-10-775-169-2172	Sequence 2172, Ap
97	12.2	61.0	31	17	US-10-617-070-517	Sequence 517, App
98	12.2	61.0	32	15	US-10-287-919-1229	Sequence 1229, Ap
99	12.2	61.0	32	17	US-10-617-070-506	Sequence 506, App
c 100	12.2	61.0	39	17	US-10-371-913-43	Sequence 43, Appl
c 101	12.2	61.0	39	17	US-10-606-577-42	Sequence 42, Appl
c 102	12.2	61.0	41	10	US-09-964-059B-149	Sequence 149, App
c 103	12.2	61.0	41	17	US-10-430-201-1336	Sequence 1336, Ap
c 104	12.2	61.0	41	17	US-10-430-201-1337	Sequence 1337, Ap
c 105	12.2	61.0	41	17	US-10-430-201-1338	Sequence 1338, Ap
106	12.2	61.0	42	15	US-10-218-567-33	Sequence 33, Appl
c 107	12.2	61.0	44	15	US-10-218-567-32	Sequence 32, Appl
108	12.2	61.0	45	15	US-10-062-523-15	Sequence 15, Appl
109	12.2	61.0	50	16	US-10-131-827-2886	Sequence 2886, Ap
110	12.2	61.0	50	16	US-10-131-827-4340	Sequence 4340, Ap
111	12.2	61.0	50	16	US-10-131-827-6495	Sequence 6495, Ap
c 112	12.2	61.0	50	16	US-10-131-827-6885	Sequence 6885, Ap
c 113	12.2	61.0	53	17	US-10-617-070-509	Sequence 509, App
c 114	12.2	61.0	53	17	US-10-617-070-510	Sequence 510, App
c 115	12.2	61.0	54	17	US-10-617-070-520	Sequence 520, App
c 116	12.2	61.0	54	17	US-10-617-070-521	Sequence 521, App
117	12.2	61.0	59	17	US-10-371-913-69	Sequence 69, Appl
118	12.2	61.0	59	17	US-10-606-577-66	Sequence 66, Appl
c 119	12.2	61.0	60	10	US-09-908-975-9569	Sequence 9569, Ap
c 120	12.2	61.0	60	10	US-09-908-975-23000	Sequence 23000, A
121	12.2	61.0	60	17	US-10-371-913-45	Sequence 45, Appl
122	12.2	61.0	60	17	US-10-606-577-44	Sequence 44, Appl
c 123	12.2	61.0	65	10	US-09-908-975-4341	Sequence 4341, Ap
124	12.2	61.0	72	12	US-09-987-899-1319	Sequence 1319, Ap
125	12.2	61.0	75	15	US-10-210-023-44	Sequence 44, Appl
c 126	12	60.0	24	17	US-10-664-422-226	Sequence 226, App
c 127	12	60.0	24	17	US-10-664-423-226	Sequence 226, App
c 128	12	60.0	25	15	US-10-098-263B-16287	Sequence 16287, A
129	12	60.0	25	15	US-10-098-263B-31652	Sequence 31652, A
130	12	60.0	25	15	US-10-098-263B-40054	Sequence 40054, A
131	12	60.0	25	15	US-10-098-263B-55172	Sequence 55172, A
c 132	12	60.0	25	15	US-10-098-263B-74130	Sequence 74130, A
c 133	12	60.0	25	15	US-10-098-263B-77248	Sequence 77248, A
134	12	60.0	25	15	US-10-098-263B-83525	Sequence 83525, A
135	12	60.0	25	15	US-10-098-263B-83526	Sequence 83526, A
136	12	60.0	25	15	US-10-098-263B-93575	Sequence 93575, A
c 137	12	60.0	30	15	US-10-059-273-5	Sequence 5, Appli
138	12	60.0	32	9	US-09-770-693-19	Sequence 19, Appl
c 139	12	60.0	35	10	US-09-741-744A-146	Sequence 146, App
140	12	60.0	42	16	US-10-339-268A-37	Sequence 37, Appl
141	12	60.0	42	16	US-10-339-268A-47	Sequence 47, Appl
142	12	60.0	47	16	US-10-349-143-2712	Sequence 2712, Ap
143	12	60.0	51	15	US-10-066-543-3085	Sequence 3085, Ap
144	12	60.0	55	12	US-10-035-833A-3842	Sequence 3842, Ap
145	12	60.0	57	10	US-09-918-156-61	Sequence 61, Appl
c 146	12	60.0	58	8	US-08-781-986A-4982	Sequence 4982, Ap
c 147	12	60.0	58	13	US-10-329-624-4982	Sequence 4982, Ap
c 148	12	60.0	60	10	US-09-908-975-6137	Sequence 6137, Ap
149	12	60.0	60	10	US-09-908-975-10839	Sequence 10839, A
150	12	60.0	60	10	US-09-908-975-11387	Sequence 11387, A

ALIGNMENTS

```
RESULT 1
US-10-059-273-6/c
; Sequence 6, Application US/10059273
; Publication No. US20030170736A1
; GENERAL INFORMATION:
;  APPLICANT: Agoston, Denes V.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR PRODUCING
;  TITLE OF INVENTION:  NEURAL PROGENITOR CELLS
;  FILE REFERENCE: 268422000100
;  CURRENT APPLICATION NUMBER: US/10/059,273
;  CURRENT FILING DATE:  2002-01-31
;  PRIOR APPLICATION NUMBER: US 60/265,113
;  PRIOR FILING DATE: 2001-01-31
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 37
;   TYPE: DNA
;   ORGANISM: Rat
US-10-059-273-6

  Query Match             76.0%;  Score 15.2;  DB 15;  Length 37;
  Best Local Similarity   85.0%;  Pred. No. 2.2e+03;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGTT 20
              ||| |||| || ||||||||
Db         25 AGTTACATTTACGTTTGGTT 6


RESULT 2
US-10-098-263B-60525
; Sequence 60525, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
;  APPLICANT: Mittman, Michael
;  TITLE OF INVENTION: Human Microarray
;  FILE REFERENCE: 3118.1
;  CURRENT APPLICATION NUMBER: US/10/098,263B
;  CURRENT FILING DATE:  2003-01-08
;  PRIOR APPLICATION NUMBER: 60/276,759
;  PRIOR FILING DATE: 2001-03-16
;  NUMBER OF SEQ ID NOS: 131066
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 60525
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-098-263B-60525

  Query Match             72.0%;  Score 14.4;  DB 15;  Length 25;
  Best Local Similarity   93.8%;  Pred. No. 4.8e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTT 16
              |||||||||  ||||||
Db          3 AGTAACATCGATGTTT 18


RESULT 3
US-09-842-776A-46
; Sequence 46, Application US/09842776A
; Publication No. US20040023316A1
; GENERAL INFORMATION:
;  APPLICANT: CONNEX GMBH
;  TITLE OF INVENTION: NEW METHOD FOR DETECTING ACID-RESISTANT MICROORGANISMS
;  TITLE OF INVENTION:  IN THE STOOL
;  FILE REFERENCE: 41735
;  CURRENT APPLICATION NUMBER: US/09/842,776A
;  CURRENT FILING DATE:  2002-08-15
;  PRIOR APPLICATION NUMBER: PCT/EP99/08212
;  PRIOR FILING DATE: 1999-10-29
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
;   LENGTH: 33
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
```

```
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: DNA encoding
;   OTHER INFORMATION: complementarity determining region (CDR1) of an
;   OTHER INFORMATION: antibody light chain directed to a beta-urease
;   OTHER INFORMATION: epitope (alternative sequence)
US-09-842-776A-46

  Query Match             69.0%;  Score 13.8;  DB 13;  Length 33;
  Best Local Similarity   88.2%;  Pred. No. 9.4e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 AACATCTATGTTTGGTT 20
              |||||  ||||||||||
Db         13 AACATTAATGTTTGGTT 29


RESULT 4
US-10-428-339-5/c
; Sequence 5, Application US/10428339
; Publication No. US20030228612A1
; GENERAL INFORMATION:
;  APPLICANT: KENWARD, Kimberly D.
;  APPLICANT:  SALEHUZZAMAN, Shah
;  TITLE OF INVENTION: PRODUCTION OF RECOMBINANT EPIDERMAL
;  TITLE OF INVENTION:  GROWTH FACTOR IN PLANTS
;  FILE REFERENCE: 07121.0005U2
;  CURRENT APPLICATION NUMBER: US/10/428,339
;  CURRENT FILING DATE:  2003-04-30
;  PRIOR APPLICATION NUMBER: 60/377,294
;  PRIOR FILING DATE: 2002-04-30
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 77
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence; note =
;   OTHER INFORMATION: synthetic construct
US-10-428-339-5

  Query Match             69.0%;  Score 13.8;  DB 16;  Length 77;
  Best Local Similarity   88.2%;  Pred. No. 1.1e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTG 17
              ||| |||| ||||||||
Db         55 AGTTACATGTATGTTTG 39


RESULT 5
US-10-688-706-2763/c
; Sequence 2763, Application US/10688706
; Publication No. US20040102412A1
; GENERAL INFORMATION:
;  APPLICANT: Pharmacia Corp.
;  APPLICANT:  Broschat, Kay
;  TITLE OF INVENTION: ANTISENSE MODULATION OF GFAT EXPRESSION
;  FILE REFERENCE: 01393/1
;  CURRENT APPLICATION NUMBER: US/10/688,706
;  CURRENT FILING DATE:  2003-10-17
;  PRIOR APPLICATION NUMBER: 60/419,268
;  PRIOR FILING DATE: 2002-10-17
;  NUMBER OF SEQ ID NOS: 3071
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2763
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: artificial
;   FEATURE:
;   OTHER INFORMATION: human GFAT antisense
US-10-688-706-2763

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGTT 20
              | | | ||||| ||||||||
Db         20 ATTTATATCTAAGTTTGGTT 1


RESULT 6
US-10-043-639A-3/c
; Sequence 3, Application US/10043639A
; Publication No. US20030175916A1
; GENERAL INFORMATION:
;  APPLICANT: SARCABAL, PATRICIA
;  APPLICANT:  CROUX, CHRISTIAN
;  APPLICANT:  SOUCAILLE, PHILIPPE
;  TITLE OF INVENTION: METHOD FOR PREPARING 1,3-PROPANEDIOL BY A RECOMBINANT
;  TITLE OF INVENTION:  MICRO-ORGANISM IN THE ABSENCE OF COENZYME B12 OR ONE OF
;  TITLE OF INVENTION:  ITS PRECURSORS
;  FILE REFERENCE: CHEP:004US
;  CURRENT APPLICATION NUMBER: US/10/043,639A
;  CURRENT FILING DATE:  2003-04-12
;  PRIOR APPLICATION NUMBER: PCT/FR00/01981
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: FR 99/08939
;  PRIOR FILING DATE: 1999-07-09
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 28
;   TYPE: DNA
;   ORGANISM: Clostridium butyricum
US-10-043-639A-3

  Query Match             68.0%;  Score 13.6;  DB 15;  Length 28;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGTT 20
              | |||||| | |||||| ||
Db         26 AATAACATTTTTGTTTGTTT 7


RESULT 7
US-09-911-132A-27/c
; Sequence 27, Application US/09911132A
; Publication No. US20030096341A1
; GENERAL INFORMATION:
;  APPLICANT: Roche Diagnostics GmbH
;  TITLE OF INVENTION: Expression of Alkaline Phosphatase in Yeast
;  FILE REFERENCE: RDID 0073US
;  CURRENT APPLICATION NUMBER: US/09/911,132A
;  CURRENT FILING DATE:  2002-08-28
;  NUMBER OF SEQ ID NOS: 38
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
;   LENGTH: 73
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Primer
US-09-911-132A-27

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 73;
  Best Local Similarity   80.0%;  Pred. No. 1.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGTT 20
              ||  ||  ||||| |||||||
Db         23 AGGTACTTCTATTTTTGGTT 4
```

```
RESULT 8
US-10-098-263B-17106/c
; Sequence 17106, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
;  APPLICANT: Mittman, Michael
;  TITLE OF INVENTION: Human Microarray
;  FILE REFERENCE: 3118.1
;  CURRENT APPLICATION NUMBER: US/10/098,263B
;  CURRENT FILING DATE:  2003-01-08
;  PRIOR APPLICATION NUMBER: 60/276,759
;  PRIOR FILING DATE: 2001-03-16
;  NUMBER OF SEQ ID NOS: 131066
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 17106
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-098-263B-17106

  Query Match             67.0%;  Score 13.4;  DB 15;  Length 25;
  Best Local Similarity   93.3%;  Pred. No. 1.4e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 CATCTATGTTTGGTT 20
              ||||| |||||||||
Db         16 CATCTCTGTTTGGTT 2


RESULT 9
US-09-969-373-3917/c
; Sequence 3917, Application US/09969373
; Patent No. US20020133852A1
; GENERAL INFORMATION:
;  APPLICANT: Effertz, Roger J.
;  APPLICANT:  Hauge, Brian M.
;  TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping
;  FILE REFERENCE: 38-10(52679)A
;  CURRENT APPLICATION NUMBER: US/09/969,373
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: US 09/754,853
;  PRIOR FILING DATE: 2001-01-05
;  PRIOR APPLICATION NUMBER: US 09/760,427
;  PRIOR FILING DATE: 2001-01-13
;  PRIOR APPLICATION NUMBER: US 09/855,768
;  PRIOR FILING DATE: 2001-05-15
;  NUMBER OF SEQ ID NOS: 4593
; SEQ ID NO 3917
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Glycine max
US-09-969-373-3917

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 1.7e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 TAACATCTATGTTTGGTT 20
              | | |||| |||||||||
Db         19 TCAAATCTTTGTTTGGTT 2


RESULT 10
US-10-688-706-2863/c
; Sequence 2863, Application US/10688706
; Publication No. US20040102412A1
; GENERAL INFORMATION:
;  APPLICANT: Pharmacia Corp.
;  APPLICANT:  Broschat, Kay
;  TITLE OF INVENTION: ANTISENSE MODULATION OF GFAT EXPRESSION
;  FILE REFERENCE: 01393/1
;  CURRENT APPLICATION NUMBER: US/10/688,706
;  CURRENT FILING DATE:  2003-10-17
;  PRIOR APPLICATION NUMBER: 60/419,268
;  PRIOR FILING DATE: 2002-10-17
;  NUMBER OF SEQ ID NOS: 3071
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2863
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: artificial
;   FEATURE:
;   OTHER INFORMATION: human GFAT antisense
US-10-688-706-2863

  Query Match             66.0%;  Score 13.2;  DB 17;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 1.7e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 TAACATCTATGTTTGGTT 20
              | | ||||| ||||||||
Db         19 TTATATCTAAGTTTGGTT 2


RESULT 11
US-10-688-706-2868/c
; Sequence 2868, Application US/10688706
; Publication No. US20040102412A1
; GENERAL INFORMATION:
;  APPLICANT: Pharmacia Corp.
;  APPLICANT:  Broschat, Kay
;  TITLE OF INVENTION: ANTISENSE MODULATION OF GFAT EXPRESSION
;  FILE REFERENCE: 01393/1
;  CURRENT APPLICATION NUMBER: US/10/688,706
;  CURRENT FILING DATE:  2003-10-17
;  PRIOR APPLICATION NUMBER: 60/419,268
;  PRIOR FILING DATE: 2002-10-17
;  NUMBER OF SEQ ID NOS: 3071
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2868
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: artificial
;   FEATURE:
;   OTHER INFORMATION: human GFAT antisense
US-10-688-706-2868

  Query Match             66.0%;  Score 13.2;  DB 17;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 1.7e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 TAACATCTATGTTTGGTT 20
              | | ||||| ||||||||
Db         20 TTATATCTAAGTTTGGTT 3


RESULT 12
US-09-753-143-14
; Sequence 14, Application US/09753143
; Patent No. US20020102550A1
;   GENERAL INFORMATION:
;        APPLICANT: NATHAN A. ELLIS, JAMES GERMAN, AND JOANNA
;                   GRODEN
;        TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND TREATMENT
;                            OF BLOOM'S SYNDROME
;        NUMBER OF SEQUENCES: 78
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: AMSTER, ROTHSTEIN & EBENSTEIN
;             STREET: 90 PARK AVENUE
;             CITY: NEW YORK
;             STATE: NEW YORK
;             COUNTRY: U.S.A.
;             ZIP: 10016
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5 INCH 1.44 Mb STORAGE DISKETTE
;             COMPUTER: IBM PC COMPATIBLE
```

```
;              OPERATING SYSTEM: MS-DOS
;              SOFTWARE: ASCII
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/753,143
;              FILING DATE: 02-Jan-2001
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 09/175,828
;              FILING DATE: 1998-10-20
;         ATTORNEY/AGENT INFORMATION:
;              NAME: ELIZABETH A. BOGOSIAN
;              REGISTRATION NUMBER: 39,911
;              REFERENCE/DOCKET NUMBER: 63475/65
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: (212) 697-5995
;              TELEFAX: (212) 286-0854 or 286-0082
;              TELEX: TWX 710-581-4766
;    INFORMATION FOR SEQ ID NO: 14:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 24
;              TYPE: NUCLEIC ACID
;              STRANDEDNESS: SINGLE
;              TOPOLOGY: LINEAR
;         MOLECULE TYPE: <Unknown>
;              DESCRIPTION: OTHER NUCLEIC ACID
;         HYPOTHETICAL: YES
;         ANTI-SENSE: NO
;         FEATURE:
;              NAME/KEY:
;              LOCATION:
;              IDENTIFICATION METHOD:
;              OTHER INFORMATION:
;         SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-753-143-14

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 24;
  Best Local Similarity   83.3%;  Pred. No. 1.7e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGG 18
              |||| |||| ||| ||||
Db          6 AGTACCATCAATGATTGG 23


RESULT 13
US-10-032-585-3574
; Sequence 3574, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
;  APPLICANT: Terry, Roemer D.
;  APPLICANT:  Bo, Jiang
;  APPLICANT:  Charles, Boone
;  APPLICANT:  Howard, Bussey
;  TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
;  FILE REFERENCE: 10182-005-999
;  CURRENT APPLICATION NUMBER: US/10/032,585
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 8000
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3574
;   LENGTH: 65
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-10-032-585-3574

  Query Match             66.0%;  Score 13.2;  DB 15;  Length 65;
  Best Local Similarity   83.3%;  Pred. No. 2e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGT 19
              |||||||  | |||||||
Db          7 GTAACATTCAAGTTTGGT 24


RESULT 14
US-09-877-478-145
; Sequence 145, Application US/09877478
; Publication No. US20030068301A1
; GENERAL INFORMATION:
;  APPLICANT: Ribozyme Pharmaceuticals, Inc.
;  APPLICANT:  Draper, Kenneth
;  APPLICANT:  Blatt, Larry
;  APPLICANT:  McSwiggen, Jim
;  APPLICANT:  Morrissey, Dave
;  TITLE OF INVENTION: Method and Reagent for Inhibiting Hepatitis B Virus Replication
;  FILE REFERENCE: MBHB00-845-H (400/029)
;  CURRENT APPLICATION NUMBER: US/09/877,478
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: US 07/882,712
;  PRIOR FILING DATE: 1992-05-14
;  PRIOR APPLICATION NUMBER: US 09/531,025
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: US 09/636,385
;  PRIOR FILING DATE: 2000-08-09
;  PRIOR APPLICATION NUMBER: US 09/696,347
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 08/193,627
;  PRIOR FILING DATE: 1994-02-07
;  PRIOR APPLICATION NUMBER: US 08/433,993
;  PRIOR FILING DATE: 1995-05-04
;  PRIOR APPLICATION NUMBER: US 08/434,504
;  PRIOR FILING DATE: 1995-05-04
;  PRIOR APPLICATION NUMBER: US 09/436,430
;  PRIOR FILING DATE: 1999-11-08
;  NUMBER OF SEQ ID NOS: 6586
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 145
;   LENGTH: 17
;   TYPE: RNA
;   ORGANISM: Hepatitis B virus
US-09-877-478-145

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 17;
  Best Local Similarity   50.0%;  Pred. No. 2.5e+04;
  Matches    8;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTT 16
              || ||| :|:|:|::: 
Db          1 AGGAACCUCUAUGUUU 16


RESULT 15
US-09-877-478-848
; Sequence 848, Application US/09877478
; Publication No. US20030068301A1
; GENERAL INFORMATION:
;  APPLICANT: Ribozyme Pharmaceuticals, Inc.
;  APPLICANT:  Draper, Kenneth
;  APPLICANT:  Blatt, Larry
;  APPLICANT:  McSwiggen, Jim
;  APPLICANT:  Morrissey, Dave
;  TITLE OF INVENTION: Method and Reagent for Inhibiting Hepatitis B Virus Replication
;  FILE REFERENCE: MBHB00-845-H (400/029)
;  CURRENT APPLICATION NUMBER: US/09/877,478
;  CURRENT FILING DATE:  2001-12-31
;  PRIOR APPLICATION NUMBER: US 07/882,712
;  PRIOR FILING DATE: 1992-05-14
;  PRIOR APPLICATION NUMBER: US 09/531,025
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: US 09/636,385
;  PRIOR FILING DATE: 2000-08-09
;  PRIOR APPLICATION NUMBER: US 09/696,347
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 08/193,627
;  PRIOR FILING DATE: 1994-02-07
;  PRIOR APPLICATION NUMBER: US 08/433,993
;  PRIOR FILING DATE: 1995-05-04
```

```
;  PRIOR APPLICATION NUMBER: US 08/434,504
;  PRIOR FILING DATE: 1995-05-04
;  PRIOR APPLICATION NUMBER: US 09/436,430
;  PRIOR FILING DATE: 1999-11-08
;  NUMBER OF SEQ ID NOS: 6586
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 848
;   LENGTH: 17
;   TYPE: RNA
;   ORGANISM: Hepatitis B virus
US-09-877-478-848

  Query Match           64.0%;  Score 12.8;  DB 10;  Length 17;
  Best Local Similarity   50.0%;  Pred. No. 2.5e+04;
  Matches    8;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTT 16
              || ||| :|:|:|:::
Db          2 AGGAACCUCUAUGUUU 17


RESULT 16
US-10-342-902-145
; Sequence 145, Application US/10342902
; Publication No. US20040054156A1
; GENERAL INFORMATION:
;  APPLICANT: Sirna Therapeutics, Inc.
;  APPLICANT:  Draper, Kenneth
;  APPLICANT:  Blatt, Larry
;  APPLICANT:  McSwiggen, Jim
;  APPLICANT:  Morrissey, Dave
;  TITLE OF INVENTION: Method and Reagent for Inhibiting Hepatitis B Virus Replication
;  FILE REFERENCE: 400/075 (MBHB00-845-I)
;  CURRENT APPLICATION NUMBER: US/10/342,902
;  CURRENT FILING DATE:  2003-01-15
;  PRIOR APPLICATION NUMBER: US 09/877,478
;  PRIOR FILING DATE: 2001-06-08
;  PRIOR APPLICATION NUMBER: US 09/531,025
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: US 09/636,385
;  PRIOR FILING DATE: 2000-08-09
;  PRIOR APPLICATION NUMBER: US 09/696,347
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 08/193,627
;  PRIOR FILING DATE: 1994-02-07
;  PRIOR APPLICATION NUMBER: US 07/882,712
;  PRIOR FILING DATE: 1992-05-14
;  PRIOR APPLICATION NUMBER: US 09/436,430
;  PRIOR FILING DATE: 1999-11-08
;  NUMBER OF SEQ ID NOS: 6592
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 145
;   LENGTH: 17
;   TYPE: RNA
;   ORGANISM: Hepatitis B virus
US-10-342-902-145

  Query Match           64.0%;  Score 12.8;  DB 13;  Length 17;
  Best Local Similarity   50.0%;  Pred. No. 2.5e+04;
  Matches    8;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTT 16
              || ||| :|:|:|:::
Db          1 AGGAACCUCUAUGUUU 16


RESULT 17
US-10-342-902-848
; Sequence 848, Application US/10342902
; Publication No. US20040054156A1
; GENERAL INFORMATION:
;  APPLICANT: Sirna Therapeutics, Inc.
;  APPLICANT:  Draper, Kenneth
;  APPLICANT:  Blatt, Larry
;  APPLICANT:  McSwiggen, Jim
;  APPLICANT:  Morrissey, Dave
;  TITLE OF INVENTION: Method and Reagent for Inhibiting Hepatitis B Virus Replication
;  FILE REFERENCE: 400/075 (MBHB00-845-I)
;  CURRENT APPLICATION NUMBER: US/10/342,902
;  CURRENT FILING DATE:  2003-01-15
;  PRIOR APPLICATION NUMBER: US 09/877,478
;  PRIOR FILING DATE: 2001-06-08
;  PRIOR APPLICATION NUMBER: US 09/531,025
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: US 09/636,385
;  PRIOR FILING DATE: 2000-08-09
;  PRIOR APPLICATION NUMBER: US 09/696,347
;  PRIOR FILING DATE: 2000-10-24
;  PRIOR APPLICATION NUMBER: US 08/193,627
;  PRIOR FILING DATE: 1994-02-07
;  PRIOR APPLICATION NUMBER: US 07/882,712
;  PRIOR FILING DATE: 1992-05-14
;  PRIOR APPLICATION NUMBER: US 09/436,430
;  PRIOR FILING DATE: 1999-11-08
;  NUMBER OF SEQ ID NOS: 6592
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 848
;   LENGTH: 17
;   TYPE: RNA
;   ORGANISM: Hepatitis B virus
US-10-342-902-848

  Query Match           64.0%;  Score 12.8;  DB 13;  Length 17;
  Best Local Similarity   50.0%;  Pred. No. 2.5e+04;
  Matches    8;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTT 16
              || ||| :|:|:|:::
Db          2 AGGAACCUCUAUGUUU 17


RESULT 18
US-10-669-841-145
; Sequence 145, Application US/10669841
; Publication No. US20040127446A1
; GENERAL INFORMATION:
;  APPLICANT: Sirna Therapeutics, Inc.
;  APPLICANT:  Lawrence, Blatt
;  APPLICANT:  Dennis, Macejak
;  APPLICANT:  James, McSwiggen
;  APPLICANT:  David, Morrissey
;  APPLICANT:  Pamela, Pavco
;  APPLICANT:  Patrice, Lee
;  APPLICANT:  Kenneth, Draper
;  APPLICANT:  Elisabeth, Roberts
;  TITLE OF INVENTION: OLIGONUCLEOTIDE MEDIATED INHIBITION OF HEPATITIS B VIRUS AND HEPA
;  TITLE OF INVENTION:  VIRUS REPLICATION
;  FILE REFERENCE: 400/042US  (MBHB02-249-E)
;  CURRENT APPLICATION NUMBER: US/10/669,841
;  CURRENT FILING DATE:  2003-09-23
;  PRIOR APPLICATION NUMBER: PCT/US02/09187
;  PRIOR FILING DATE: 2002-03-26
;  PRIOR APPLICATION NUMBER: US 60/296,876
;  PRIOR FILING DATE: 2001-06-08
;  PRIOR APPLICATION NUMBER: US 60/335,059
;  PRIOR FILING DATE: 2001-10-24
;  PRIOR APPLICATION NUMBER: US 60/337,055
;  PRIOR FILING DATE: 2001-12-05
;  PRIOR APPLICATION NUMBER: US 60/358,580
;  PRIOR FILING DATE: 2002-02-20
;  PRIOR APPLICATION NUMBER: US 60/363,124
;  PRIOR FILING DATE: 2002-03-11
;  PRIOR APPLICATION NUMBER: US 09/817,879
;  PRIOR FILING DATE: 2001-03-26
;  PRIOR APPLICATION NUMBER: US 09/740,332
;  PRIOR FILING DATE: 2000-12-18
```

```
;  PRIOR APPLICATION NUMBER: US 09/611,931
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: US 09/504,321
;  PRIOR FILING DATE: 2000-02-15
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 16207
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 145
;   LENGTH: 17
;   TYPE: RNA
;   ORGANISM: Hepatitis B Virus
US-10-669-841-145

  Query Match             64.0%;  Score 12.8;  DB 17;  Length 17;
  Best Local Similarity   50.0%;  Pred. No. 2.5e+04;
  Matches    8;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTT 16
              || ||| :|:|:|:::
Db          1 AGGAACCUCUAUGUUU 16


RESULT 19
US-10-669-841-848
; Sequence 848, Application US/10669841
; Publication No. US20040127446A1
; GENERAL INFORMATION:
;  APPLICANT: Sirna Therapeutics, Inc.
;  APPLICANT:  Lawrence, Blatt
;  APPLICANT:  Dennis, Macejak
;  APPLICANT:  James, McSwiggen
;  APPLICANT:  David, Morrissey
;  APPLICANT:  Pamela, Pavco
;  APPLICANT:  Patrice, Lee
;  APPLICANT:  Kenneth, Draper
;  APPLICANT:  Elisabeth, Roberts
;  TITLE OF INVENTION: OLIGONUCLEOTIDE MEDIATED INHIBITION OF HEPATITIS B VIRUS AND HEPA
;  TITLE OF INVENTION:  VIRUS REPLICATION
;  FILE REFERENCE: 400/042US  (MBHB02-249-E)
;  CURRENT APPLICATION NUMBER: US/10/669,841
;  CURRENT FILING DATE:  2003-09-23
;  PRIOR APPLICATION NUMBER: PCT/US02/09187
;  PRIOR FILING DATE: 2002-03-26
;  PRIOR APPLICATION NUMBER: US 60/296,876
;  PRIOR FILING DATE: 2001-06-08
;  PRIOR APPLICATION NUMBER: US 60/335,059
;  PRIOR FILING DATE: 2001-10-24
;  PRIOR APPLICATION NUMBER: US 60/337,055
;  PRIOR FILING DATE: 2001-12-05
;  PRIOR APPLICATION NUMBER: US 60/358,580
;  PRIOR FILING DATE: 2002-02-20
;  PRIOR APPLICATION NUMBER: US 60/363,124
;  PRIOR FILING DATE: 2002-03-11
;  PRIOR APPLICATION NUMBER: US 09/817,879
;  PRIOR FILING DATE: 2001-03-26
;  PRIOR APPLICATION NUMBER: US 09/740,332
;  PRIOR FILING DATE: 2000-12-18
;  PRIOR APPLICATION NUMBER: US 09/611,931
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: US 09/504,321
;  PRIOR FILING DATE: 2000-02-15
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 16207
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 848
;   LENGTH: 17
;   TYPE: RNA
;   ORGANISM: Hepatitis B Virus
US-10-669-841-848

  Query Match             64.0%;  Score 12.8;  DB 17;  Length 17;
  Best Local Similarity   50.0%;  Pred. No. 2.5e+04;
  Matches    8;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTT 16
              || ||| :|:|:|:::
Db          2 AGGAACCUCUAUGUUU 17


RESULT 20
US-10-688-706-3025/c
; Sequence 3025, Application US/10688706
; Publication No. US20040102412A1
; GENERAL INFORMATION:
;  APPLICANT: Pharmacia Corp.
;  APPLICANT:  Broschat, Kay
;  TITLE OF INVENTION: ANTISENSE MODULATION OF GFAT EXPRESSION
;  FILE REFERENCE: 01393/1
;  CURRENT APPLICATION NUMBER: US/10/688,706
;  CURRENT FILING DATE:  2003-10-17
;  PRIOR APPLICATION NUMBER: 60/419,268
;  PRIOR FILING DATE: 2002-10-17
;  NUMBER OF SEQ ID NOS: 3071
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3025
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: artificial
;   FEATURE:
;   OTHER INFORMATION: human GFAT antisense
US-10-688-706-3025

  Query Match             64.0%;  Score 12.8;  DB 17;  Length 20;
  Best Local Similarity   87.5%;  Pred. No. 2.5e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGGTT 20
              | ||||| ||||||||
Db         19 ATATCTAAGTTTGGTT 4


RESULT 21
US-10-688-706-3034/c
; Sequence 3034, Application US/10688706
; Publication No. US20040102412A1
; GENERAL INFORMATION:
;  APPLICANT: Pharmacia Corp.
;  APPLICANT:  Broschat, Kay
;  TITLE OF INVENTION: ANTISENSE MODULATION OF GFAT EXPRESSION
;  FILE REFERENCE: 01393/1
;  CURRENT APPLICATION NUMBER: US/10/688,706
;  CURRENT FILING DATE:  2003-10-17
;  PRIOR APPLICATION NUMBER: 60/419,268
;  PRIOR FILING DATE: 2002-10-17
;  NUMBER OF SEQ ID NOS: 3071
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3034
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: artificial
;   FEATURE:
;   OTHER INFORMATION: human GFAT antisense
US-10-688-706-3034

  Query Match             64.0%;  Score 12.8;  DB 17;  Length 20;
  Best Local Similarity   87.5%;  Pred. No. 2.5e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGGTT 20
              | ||||| ||||||||
Db         20 ATATCTAAGTTTGGTT 5


RESULT 22
US-10-215-112-6520
; Sequence 6520, Application US/10215112
```

```
; Publication No. US20030082596A1
; GENERAL INFORMATION:
;  APPLICANT: Michael Mittmann
;  TITLE OF INVENTION: Method of Genetic Analysis of Probes:
;  TITLE OF INVENTION:  Test3
;  FILE REFERENCE: 3119
;  CURRENT APPLICATION NUMBER: US/10/215,112
;  CURRENT FILING DATE:  2002-08-08
;  NUMBER OF SEQ ID NOS: 14936
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6520
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-6520

  Query Match             64.0%;  Score 12.8;  DB 15;  Length 25;
  Best Local Similarity   87.5%;  Pred. No. 2.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 AACATCTATGTTTGGT 19
              ||| |||||||||| |
Db          2 AACTTCTATGTTTGCT 17


RESULT 23
US-10-215-112-10981
; Sequence 10981, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
;  APPLICANT: Michael Mittmann
;  TITLE OF INVENTION: Method of Genetic Analysis of Probes:
;  TITLE OF INVENTION:  Test3
;  FILE REFERENCE: 3119
;  CURRENT APPLICATION NUMBER: US/10/215,112
;  CURRENT FILING DATE:  2002-08-08
;  NUMBER OF SEQ ID NOS: 14936
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10981
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-10981

  Query Match             64.0%;  Score 12.8;  DB 15;  Length 25;
  Best Local Similarity   87.5%;  Pred. No. 2.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 AACATCTATGTTTGGT 19
              ||| |||||||||| |
Db          4 AACTTCTATGTTTGCT 19


RESULT 24
US-10-098-263B-1523
; Sequence 1523, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
;  APPLICANT: Mittman, Michael
;  TITLE OF INVENTION: Human Microarray
;  FILE REFERENCE: 3118.1
;  CURRENT APPLICATION NUMBER: US/10/098,263B
;  CURRENT FILING DATE:  2003-01-08
;  PRIOR APPLICATION NUMBER: 60/276,759
;  PRIOR FILING DATE: 2001-03-16
;  NUMBER OF SEQ ID NOS: 131066
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 1523
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-098-263B-1523

  Query Match             64.0%;  Score 12.8;  DB 15;  Length 25;
  Best Local Similarity   87.5%;  Pred. No. 2.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 AACATCTATGTTTGGT 19
              ||| ||||| ||||||
Db          1 AACGTCTATCTTTGGT 16


RESULT 25
US-10-098-263B-60526
; Sequence 60526, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
;  APPLICANT: Mittman, Michael
;  TITLE OF INVENTION: Human Microarray
;  FILE REFERENCE: 3118.1
;  CURRENT APPLICATION NUMBER: US/10/098,263B
;  CURRENT FILING DATE:  2003-01-08
;  PRIOR APPLICATION NUMBER: 60/276,759
;  PRIOR FILING DATE: 2001-03-16
;  NUMBER OF SEQ ID NOS: 131066
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 60526
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-098-263B-60526

  Query Match             64.0%;  Score 12.8;  DB 15;  Length 25;
  Best Local Similarity   87.5%;  Pred. No. 2.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTT 16
              |||||||||  |||||
Db          3 AGTAACATCGTTGTTT 18


RESULT 26
US-10-775-169-3303/c
; Sequence 3303, Application US/10775169
; Publication No. US20040175743A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Burczynski, Michael
;  APPLICANT:  Twine, Natalie
;  APPLICANT:  Dorner, Andrew
;  APPLICANT:  Trepicchio, William
;  TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
;  FILE REFERENCE: AM101080 (031896-013000)
;  CURRENT APPLICATION NUMBER: US/10/775,169
;  CURRENT FILING DATE:  2004-02-11
;  NUMBER OF SEQ ID NOS: 5278
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3303
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: probe
US-10-775-169-3303

  Query Match             64.0%;  Score 12.8;  DB 17;  Length 25;
  Best Local Similarity   87.5%;  Pred. No. 2.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 AACATCTATGTTTGGT 19
              | ||| ||||||||||
Db         23 AGCATTTATGTTTGGT 8
```

```
RESULT 27
US-10-775-169-3304/c
; Sequence 3304, Application US/10775169
; Publication No. US20040175743A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Burczynski, Michael
;  APPLICANT:  Twine, Natalie
;  APPLICANT:  Dorner, Andrew
;  APPLICANT:  Trepicchio, William
;  TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
;  FILE REFERENCE: AM101080 (031896-013000)
;  CURRENT APPLICATION NUMBER: US/10/775,169
;  CURRENT FILING DATE:  2004-02-11
;  NUMBER OF SEQ ID NOS: 5278
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3304
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: probe
US-10-775-169-3304

  Query Match             64.0%;  Score 12.8;  DB 17;  Length 25;
  Best Local Similarity   87.5%;  Pred. No. 2.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 AACATCTATGTTTGGT 19
              | ||| ||||||||||
Db         20 AGCATTTATGTTTGGT 5


RESULT 28
US-10-002-623-548
; Sequence 548, Application US/10002623
; Publication No. US20030134285A1
; GENERAL INFORMATION:
;  APPLICANT: OEFNER, PETER J.
;  APPLICANT:  UNDERHILL, PETER A.
;  TITLE OF INVENTION: A METHOD FOR DETERMINING GENETIC
;  TITLE OF INVENTION:  AFFILIATION, SUBSTRUCTURE AND GENE FLOW WITHIN HUMAN
;  TITLE OF INVENTION:  POPULATIONS
;  FILE REFERENCE: STAN-212
;  CURRENT APPLICATION NUMBER: US/10/002,623
;  CURRENT FILING DATE:  2001-11-01
;  PRIOR APPLICATION NUMBER: US 60/245,355
;  PRIOR FILING DATE: 2000-11-01
;  NUMBER OF SEQ ID NOS: 952
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 548
;   LENGTH: 27
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-10-002-623-548

  Query Match             64.0%;  Score 12.8;  DB 15;  Length 27;
  Best Local Similarity   87.5%;  Pred. No. 2.7e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTT 16
              |||||||||||| | |
Db          7 AGTAACATCTATTTCT 22


RESULT 29
US-10-168-445-216
; Sequence 216, Application US/10168445
; Publication No. US20030177518A1
; GENERAL INFORMATION:
;  APPLICANT: Osbourn, Anne E
;  APPLICANT:  Haralampidis, Kosmas
;  APPLICANT:  Bryan, Gregory T
;  TITLE OF INVENTION: Plant Gene
;  FILE REFERENCE: 0380-P02892US0
;  CURRENT APPLICATION NUMBER: US/10/168,445
;  CURRENT FILING DATE:  2002-10-30
;  PRIOR APPLICATION NUMBER: PCT/GB00/04908
;  PRIOR FILING DATE: 2000-12-20
;  PRIOR APPLICATION NUMBER: GB 9930394.3
;  PRIOR FILING DATE: 1999-12-22
;  PRIOR APPLICATION NUMBER: GB 0020217.6
;  PRIOR FILING DATE: 2000-08-16
;  NUMBER OF SEQ ID NOS: 219
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 216
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: Primer
US-10-168-445-216

  Query Match             64.0%;  Score 12.8;  DB 15;  Length 30;
  Best Local Similarity   87.5%;  Pred. No. 2.7e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGGTT 20
              ||||| ||||||| ||
Db          7 ACATCCATGTTTGTTT 22


RESULT 30
US-10-353-274-3
; Sequence 3, Application US/10353274
; Publication No. US20030235899A1
; GENERAL INFORMATION:
;  APPLICANT: Agouron Pharmaceuticals, Inc.
;  TITLE OF INVENTION: CATALYTIC DOMAIN OF THE HUMAN EFFECTOR CELL CYCLE CHECKPOINT PROT
;  TITLE OF INVENTION:  KINASE, CHK1, MATERIALS AND METHODS FOR IDENTIFICATION OF INHIBI
;  FILE REFERENCE: PC19060B
;  CURRENT APPLICATION NUMBER: US/10/353,274
;  CURRENT FILING DATE:  2003-01-28
;  PRIOR APPLICATION NUMBER: US 09/460,421
;  PRIOR FILING DATE: 1999-12-14
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 34
;   TYPE: DNA
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Seqeunce: PCR primer
US-10-353-274-3

  Query Match             64.0%;  Score 12.8;  DB 16;  Length 34;
  Best Local Similarity   87.5%;  Pred. No. 2.8e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTT 16
              |||| ||||||| |||
Db          7 AGTACCATCTATCTTT 22


RESULT 31
US-10-027-632-58587/c
; Sequence 58587, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
```

```
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 58587
;   LENGTH: 36
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-58587

  Query Match             64.0%;  Score 12.8;  DB 13;  Length 36;
  Best Local Similarity   77.8%;  Pred. No. 2.8e+04;
  Matches   14;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGT 19
              | ||| ||: ||||||||
Db         18 GGAACTTCYGTGTTTGGT 1


RESULT 32
US-10-027-632-58595/c
; Sequence 58595, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 58595
;   LENGTH: 36
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-58595

  Query Match             64.0%;  Score 12.8;  DB 13;  Length 36;
  Best Local Similarity   77.8%;  Pred. No. 2.8e+04;
  Matches   14;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGT 19
              | ||| ||: ||||||||
Db         18 GGAACTTCYGTGTTTGGT 1


RESULT 33
US-10-027-632-58587/c
; Sequence 58587, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 58587
;   LENGTH: 36
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-58587

  Query Match             64.0%;  Score 12.8;  DB 16;  Length 36;
  Best Local Similarity   77.8%;  Pred. No. 2.8e+04;
  Matches   14;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGT 19
              | ||| ||: ||||||||
Db         18 GGAACTTCYGTGTTTGGT 1


RESULT 34
US-10-027-632-58595/c
; Sequence 58595, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 58595
;   LENGTH: 36
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-58595

  Query Match             64.0%;  Score 12.8;  DB 16;  Length 36;
```

```
  Best Local Similarity   77.8%;  Pred. No. 2.8e+04;
  Matches   14;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGT 19
              | ||| ||: ||||||||
Db         18 GGAACTTCYGTGTTTGGT 1


RESULT 35
US-10-032-585-1712/c
; Sequence 1712, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
;  APPLICANT: Terry, Roemer D.
;  APPLICANT:  Bo, Jiang
;  APPLICANT:  Charles, Boone
;  APPLICANT:  Howard, Bussey
;  TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
;  FILE REFERENCE: 10182-005-999
;  CURRENT APPLICATION NUMBER: US/10/032,585
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 8000
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1712
;   LENGTH: 43
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-10-032-585-1712

  Query Match             64.0%;  Score 12.8;  DB 15;  Length 43;
  Best Local Similarity   87.5%;  Pred. No. 2.9e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGGTT 20
              | |||||||||||  ||
Db         31 AGATCTATGTTTGTTT 16


RESULT 36
US-10-349-143-1887/c
; Sequence 1887, Application US/10349143
; Publication No. US20040005584A1
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Chumakov, Ilya
;  TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
;  FILE REFERENCE: GENSET.020CP1
;  CURRENT APPLICATION NUMBER: US/10/349,143
;  CURRENT FILING DATE:  2003-01-21
;  PRIOR APPLICATION NUMBER: US/09/422,978
;  PRIOR FILING DATE: 1999-10-20
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/298,850
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-21
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/109,732
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-23
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/082,614
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-21
;  NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 1887
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: 99-7129-335 : polymorphic base A or C
US-10-349-143-1887

  Query Match             64.0%;  Score 12.8;  DB 16;  Length 47;
  Best Local Similarity   87.5%;  Pred. No. 2.9e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGGTT 20
              |||| ||||||||  ||
Db         43 ACATTTATGTTTGTTT 28


RESULT 37
US-10-688-706-2716/c
; Sequence 2716, Application US/10688706
; Publication No. US20040102412A1
; GENERAL INFORMATION:
;  APPLICANT: Pharmacia Corp.
;  APPLICANT:  Broschat, Kay
;  TITLE OF INVENTION: ANTISENSE MODULATION OF GFAT EXPRESSION
;  FILE REFERENCE: 01393/1
;  CURRENT APPLICATION NUMBER: US/10/688,706
;  CURRENT FILING DATE:  2003-10-17
;  PRIOR APPLICATION NUMBER: 60/419,268
;  PRIOR FILING DATE: 2002-10-17
;  NUMBER OF SEQ ID NOS: 3071
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2716
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: artificial
;   FEATURE:
;   OTHER INFORMATION: human GFAT antisense
US-10-688-706-2716

  Query Match             63.0%;  Score 12.6;  DB 17;  Length 20;
  Best Local Similarity   78.9%;  Pred. No. 3.1e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              | | | ||||| |||||||
Db         19 ATTTATATCTAAGTTTGGT 1


RESULT 38
US-10-098-263B-25123/c
; Sequence 25123, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
;  APPLICANT: Mittman, Michael
;  TITLE OF INVENTION: Human Microarray
;  FILE REFERENCE: 3118.1
;  CURRENT APPLICATION NUMBER: US/10/098,263B
;  CURRENT FILING DATE:  2003-01-08
;  PRIOR APPLICATION NUMBER: 60/276,759
;  PRIOR FILING DATE: 2001-03-16
;  NUMBER OF SEQ ID NOS: 131066
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 25123
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-098-263B-25123

  Query Match             63.0%;  Score 12.6;  DB 15;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 3.2e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGTT 20
              | |  |||||| |||||||
Db         23 GAAGGATCTATCTTTGGTT 5


RESULT 39
US-10-098-263B-27074
; Sequence 27074, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
;  APPLICANT: Mittman, Michael
;  TITLE OF INVENTION: Human Microarray
```

```
;  FILE REFERENCE: 3118.1
;  CURRENT APPLICATION NUMBER: US/10/098,263B
;  CURRENT FILING DATE:  2003-01-08
;  PRIOR APPLICATION NUMBER: 60/276,759
;  PRIOR FILING DATE: 2001-03-16
;  NUMBER OF SEQ ID NOS: 131066
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 27074
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-098-263B-27074

  Query Match             63.0%;  Score 12.6;  DB 15;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 3.2e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              |||||  |  ||||||| ||
Db          1 AGTAATAGATATGTTTCGT 19


RESULT 40
US-10-098-263B-29910/c
; Sequence 29910, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
;  APPLICANT: Mittman, Michael
;  TITLE OF INVENTION: Human Microarray
;  FILE REFERENCE: 3118.1
;  CURRENT APPLICATION NUMBER: US/10/098,263B
;  CURRENT FILING DATE:  2003-01-08
;  PRIOR APPLICATION NUMBER: 60/276,759
;  PRIOR FILING DATE: 2001-03-16
;  NUMBER OF SEQ ID NOS: 131066
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 29910
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-098-263B-29910

  Query Match             63.0%;  Score 12.6;  DB 15;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 3.2e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              |||||||||  | ||  || |
Db         19 AGTAACATCAAAGTCTGTT 1


RESULT 41
US-10-098-263B-35597
; Sequence 35597, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
;  APPLICANT: Mittman, Michael
;  TITLE OF INVENTION: Human Microarray
;  FILE REFERENCE: 3118.1
;  CURRENT APPLICATION NUMBER: US/10/098,263B
;  CURRENT FILING DATE:  2003-01-08
;  PRIOR APPLICATION NUMBER: 60/276,759
;  PRIOR FILING DATE: 2001-03-16
;  NUMBER OF SEQ ID NOS: 131066
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 35597
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-098-263B-35597

  Query Match             63.0%;  Score 12.6;  DB 15;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 3.2e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGTT 20
              ||||||  || |||| |||
Db          1 GTAACAGGTAGGTTTCGTT 19


RESULT 42
US-10-098-263B-35736/c
; Sequence 35736, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
;  APPLICANT: Mittman, Michael
;  TITLE OF INVENTION: Human Microarray
;  FILE REFERENCE: 3118.1
;  CURRENT APPLICATION NUMBER: US/10/098,263B
;  CURRENT FILING DATE:  2003-01-08
;  PRIOR APPLICATION NUMBER: 60/276,759
;  PRIOR FILING DATE: 2001-03-16
;  NUMBER OF SEQ ID NOS: 131066
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 35736
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-098-263B-35736

  Query Match             63.0%;  Score 12.6;  DB 15;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 3.2e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              ||||||  || ||||  || |
Db         24 AGTAACTTCAATGTCTGTT 6


RESULT 43
US-10-098-263B-36372/c
; Sequence 36372, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
;  APPLICANT: Mittman, Michael
;  TITLE OF INVENTION: Human Microarray
;  FILE REFERENCE: 3118.1
;  CURRENT APPLICATION NUMBER: US/10/098,263B
;  CURRENT FILING DATE:  2003-01-08
;  PRIOR APPLICATION NUMBER: 60/276,759
;  PRIOR FILING DATE: 2001-03-16
;  NUMBER OF SEQ ID NOS: 131066
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 36372
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-098-263B-36372

  Query Match             63.0%;  Score 12.6;  DB 15;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 3.2e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              ||||||  || ||||  || |
Db         24 AGTAACTTCAATGTCTGTT 6


RESULT 44
US-10-098-263B-65195
; Sequence 65195, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
;  APPLICANT: Mittman, Michael
;  TITLE OF INVENTION: Human Microarray
;  FILE REFERENCE: 3118.1
```

; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 65195
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-65195

Query Match 63.0%; Score 12.6; DB 15; Length 25;
Best Local Similarity 78.9%; Pred. No. 3.2e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          2 GTAACATCTATGTTTGGTT 20
              |||||| || |||| |||
Db          2 GTAACAGGTAGGTTTCGTT 20
```

RESULT 45
US-10-098-263B-100102
; Sequence 100102, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 100102
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-100102

Query Match 63.0%; Score 12.6; DB 15; Length 25;
Best Local Similarity 78.9%; Pred. No. 3.2e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 AGTAACATCTATGTTTGGT 19
              |||| |||||| |||  ||
Db          3 AGTACCATCTACGTTCCGT 21
```

Search completed: September 23, 2004, 16:48:21
Job time : 234 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 23, 2004, 15:45:09 ; Search time 53 Seconds
(without alignments)
209.415 Million cell updates/sec

Title: US-10-798-923A-36
Perfect score: 20
Sequence: 1 agtaacatctatgtttggtt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 915622

Minimum DB seq length: 0
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	13.8	69.0	20	3	US-09-359-757-34	Sequence 34, Appl
	2	13.8	69.0	43	3	US-09-187-050-9	Sequence 9, Appli
	3	13.2	66.0	24	1	US-08-559-303B-14	Sequence 14, Appl
	4	13.2	66.0	24	3	US-09-175-828-14	Sequence 14, Appl
	5	12.8	64.0	34	4	US-09-460-421-3	Sequence 3, Appli
c	6	12.8	64.0	47	4	US-09-422-978-1887	Sequence 1887, Ap
c	7	12.8	64.0	74	3	US-09-315-793-53	Sequence 53, Appl
c	8	12.6	63.0	24	3	US-08-868-699A-5	Sequence 5, Appli
c	9	12.6	63.0	24	4	US-09-757-014-5	Sequence 5, Appli
c	10	12.6	63.0	27	3	US-09-009-913-223	Sequence 223, App
	11	12.6	63.0	34	1	US-07-749-446-3	Sequence 3, Appli
	12	12.6	63.0	35	3	US-08-584-760A-50	Sequence 50, Appl
c	13	12.6	63.0	36	3	US-08-584-760A-49	Sequence 49, Appl
c	14	12.6	63.0	37	4	US-09-086-726-11	Sequence 11, Appl
c	15	12.2	61.0	34	1	US-08-233-130A-3	Sequence 3, Appli
	16	12	60.0	22	4	US-09-270-140A-52	Sequence 52, Appl
	17	12	60.0	30	2	US-08-629-001A-60	Sequence 60, Appl
	18	12	60.0	30	3	US-08-642-274D-139	Sequence 139, App
	19	12	60.0	45	1	US-08-233-009-51	Sequence 51, Appl
	20	12	60.0	47	4	US-09-422-978-2712	Sequence 2712, Ap
c	21	12	60.0	51	4	US-09-443-199C-847	Sequence 847, App
c	22	12	60.0	51	4	US-09-443-199C-848	Sequence 848, App
	23	12	60.0	57	3	US-08-864-473-61	Sequence 61, Appl
	24	12	60.0	57	3	US-09-440-523-61	Sequence 61, Appl
c	25	12	60.0	58	4	US-08-956-171E-4982	Sequence 4982, Ap
	26	12	60.0	63	1	US-08-303-275-198	Sequence 198, App
c	27	12	60.0	63	1	US-08-303-275-199	Sequence 199, App
	28	11.8	59.0	16	4	US-09-479-005A-399	Sequence 399, App
	29	11.8	59.0	18	1	US-08-261-822A-57	Sequence 57, Appl
	30	11.8	59.0	18	5	PCT-US95-07744A-57	Sequence 57, Appl
c	31	11.8	59.0	20	3	US-09-287-796-149	Sequence 149, App
c	32	11.8	59.0	20	3	US-09-130-616-149	Sequence 149, App
	33	11.8	59.0	29	1	US-08-445-746-10	Sequence 10, Appl
	34	11.8	59.0	29	3	US-09-008-722-10	Sequence 10, Appl
	35	11.8	59.0	32	1	US-08-320-373-1	Sequence 1, Appli
c	36	11.8	59.0	33	1	US-08-320-373-2	Sequence 2, Appli
c	37	11.8	59.0	50	4	US-08-956-171E-2061	Sequence 2061, Ap
	38	11.8	59.0	54	4	US-08-956-171E-5031	Sequence 5031, Ap
c	39	11.8	59.0	58	4	US-08-956-171E-4977	Sequence 4977, Ap
	40	11.8	59.0	65	4	US-09-462-941-30	Sequence 30, Appl
	41	11.8	59.0	66	4	US-09-462-941-29	Sequence 29, Appl
	42	11.8	59.0	67	1	US-08-398-617-5	Sequence 5, Appli
c	43	11.8	59.0	67	1	US-08-398-617-6	Sequence 6, Appli
	44	11.8	59.0	67	2	US-08-398-615-5	Sequence 5, Appli
c	45	11.8	59.0	67	2	US-08-398-615-6	Sequence 6, Appli
	46	11.8	59.0	67	3	US-08-397-303-5	Sequence 5, Appli
c	47	11.8	59.0	67	3	US-08-397-303-6	Sequence 6, Appli
	48	11.8	59.0	69	1	US-08-435-501-27	Sequence 27, Appl
	49	11.8	59.0	69	1	US-08-435-764-27	Sequence 27, Appl
	50	11.8	59.0	69	5	PCT-US93-04717-27	Sequence 27, Appl
	51	11.8	59.0	72	4	US-09-786-569-8	Sequence 8, Appli
c	52	11.8	59.0	72	4	US-09-786-569-9	Sequence 9, Appli
c	53	11.8	59.0	75	1	US-08-398-617-4	Sequence 4, Appli
c	54	11.8	59.0	75	2	US-08-398-615-4	Sequence 4, Appli
c	55	11.8	59.0	75	3	US-08-397-303-4	Sequence 4, Appli
	56	11.8	59.0	77	4	US-08-956-171E-4957	Sequence 4957, Ap
	57	11.8	59.0	78	1	US-08-398-617-16	Sequence 16, Appl
	58	11.8	59.0	78	1	US-08-398-617-18	Sequence 18, Appl
	59	11.8	59.0	78	1	US-08-398-617-19	Sequence 19, Appl
	60	11.8	59.0	78	1	US-08-398-617-20	Sequence 20, Appl
	61	11.8	59.0	78	1	US-08-398-617-21	Sequence 21, Appl
	62	11.8	59.0	78	1	US-08-398-617-22	Sequence 22, Appl
	63	11.8	59.0	78	2	US-08-398-615-16	Sequence 16, Appl
	64	11.8	59.0	78	2	US-08-398-615-18	Sequence 18, Appl
	65	11.8	59.0	78	2	US-08-398-615-19	Sequence 19, Appl
	66	11.8	59.0	78	2	US-08-398-615-20	Sequence 20, Appl
	67	11.8	59.0	78	2	US-08-398-615-21	Sequence 21, Appl
	68	11.8	59.0	78	2	US-08-398-615-22	Sequence 22, Appl
	69	11.8	59.0	78	3	US-08-397-303-16	Sequence 16, Appl
	70	11.8	59.0	78	3	US-08-397-303-18	Sequence 18, Appl
	71	11.8	59.0	78	3	US-08-397-303-19	Sequence 19, Appl
	72	11.8	59.0	78	3	US-08-397-303-20	Sequence 20, Appl
	73	11.8	59.0	78	3	US-08-397-303-21	Sequence 21, Appl
	74	11.8	59.0	78	3	US-08-397-303-22	Sequence 22, Appl
	75	11.8	59.0	79	1	US-08-398-617-11	Sequence 11, Appl
c	76	11.8	59.0	79	1	US-08-398-617-12	Sequence 12, Appl
	77	11.8	59.0	79	2	US-08-398-615-11	Sequence 11, Appl
c	78	11.8	59.0	79	2	US-08-398-615-12	Sequence 12, Appl
	79	11.8	59.0	79	3	US-08-397-303-11	Sequence 11, Appl
c	80	11.8	59.0	79	3	US-08-397-303-12	Sequence 12, Appl
c	81	11.6	58.0	20	4	US-09-387-341-221	Sequence 221, App
c	82	11.6	58.0	20	4	US-09-898-361-81	Sequence 81, Appl
c	83	11.6	58.0	22	1	US-08-070-328A-2	Sequence 2, Appli
c	84	11.6	58.0	22	1	US-08-455-116-2	Sequence 2, Appli
	85	11.6	58.0	24	3	US-09-253-025-20	Sequence 20, Appl
	86	11.6	58.0	24	4	US-09-734-188-20	Sequence 20, Appl
	87	11.6	58.0	32	3	US-08-654-623-25	Sequence 25, Appl
c	88	11.6	58.0	35	4	US-09-599-488B-2	Sequence 2, Appli
	89	11.6	58.0	36	3	US-08-676-818-8	Sequence 8, Appli
	90	11.6	58.0	36	4	US-09-407-549-8	Sequence 8, Appli
	91	11.6	58.0	37	1	US-08-500-222-15	Sequence 15, Appl
	92	11.6	58.0	37	1	US-08-500-125-15	Sequence 15, Appl
	93	11.6	58.0	37	2	US-07-779-704B-15	Sequence 15, Appl
	94	11.6	58.0	38	1	US-08-612-895A-20	Sequence 20, Appl
c	95	11.6	58.0	38	2	US-08-185-949B-100	Sequence 100, App
	96	11.6	58.0	38	3	US-09-093-293-20	Sequence 20, Appl
	97	11.6	58.0	38	3	US-09-459-553-20	Sequence 20, Appl
	98	11.6	58.0	38	4	US-09-882-246-20	Sequence 20, Appl
	99	11.6	58.0	38	5	PCT-US94-10562A-20	Sequence 20, Appl
	100	11.6	58.0	39	1	US-07-936-533A-6	Sequence 6, Appli

c 101	11.6	58.0	39	1	US-07-936-533A-34	Sequence 34, Appl
102	11.6	58.0	39	1	US-08-344-257A-6	Sequence 6, Appli
103	11.6	58.0	43	3	US-08-654-623-47	Sequence 47, Appl
104	11.6	58.0	47	4	US-09-422-978-862	Sequence 862, App
105	11.6	58.0	49	3	US-08-434-099A-18	Sequence 18, Appl
c 106	11.6	58.0	49	3	US-08-434-099A-19	Sequence 19, Appl
107	11.6	58.0	52	6	5252476-15	Patent No. 5252476
108	11.6	58.0	55	3	US-08-654-623-49	Sequence 49, Appl
109	11.6	58.0	57	4	US-09-428-082B-1110	Sequence 1110, Ap
c 110	11.6	58.0	57	4	US-09-428-082B-1111	Sequence 1111, Ap
111	11.6	58.0	57	4	US-09-428-082B-1113	Sequence 1113, Ap
112	11.4	57.0	17	1	US-08-373-124A-1152	Sequence 1152, Ap
113	11.4	57.0	17	1	US-08-435-628-1152	Sequence 1152, Ap
c 114	11.4	57.0	17	4	US-08-584-040-1588	Sequence 1588, Ap
c 115	11.4	57.0	17	4	US-08-584-040-1589	Sequence 1589, Ap
c 116	11.4	57.0	17	4	US-09-371-772B-133	Sequence 133, App
c 117	11.4	57.0	17	4	US-09-371-772B-134	Sequence 134, App
c 118	11.4	57.0	17	4	US-09-371-772B-4342	Sequence 4342, Ap
c 119	11.4	57.0	19	3	US-08-899-330-20	Sequence 20, Appl
c 120	11.4	57.0	20	4	US-09-668-313A-110	Sequence 110, App
121	11.4	57.0	21	3	US-09-153-555-6	Sequence 6, Appli
122	11.4	57.0	21	4	US-09-422-978-10512	Sequence 10512, A
c 123	11.4	57.0	21	4	US-08-849-303-44	Sequence 44, Appl
124	11.4	57.0	30	2	US-08-629-001A-47	Sequence 47, Appl
125	11.4	57.0	30	3	US-08-642-274D-126	Sequence 126, App
c 126	11.4	57.0	47	4	US-09-422-978-804	Sequence 804, App
c 127	11.4	57.0	47	4	US-09-422-978-1200	Sequence 1200, Ap
c 128	11.4	57.0	47	4	US-09-422-978-2933	Sequence 2933, Ap
129	11.4	57.0	70	3	US-09-275-850-139	Sequence 139, App
130	11.2	56.0	16	4	US-09-479-005A-416	Sequence 416, App
131	11.2	56.0	19	1	US-08-401-512-33	Sequence 33, Appl
132	11.2	56.0	19	3	US-09-226-012-21	Sequence 21, Appl
133	11.2	56.0	24	1	US-08-305-700-3	Sequence 3, Appli
134	11.2	56.0	24	1	US-08-528-122-3	Sequence 3, Appli
135	11.2	56.0	24	3	US-09-052-333A-68	Sequence 68, Appl
136	11.2	56.0	24	5	PCT-US95-11720-3	Sequence 3, Appli
c 137	11.2	56.0	28	1	US-08-012-543-12	Sequence 12, Appl
138	11.2	56.0	30	2	US-08-629-001A-132	Sequence 132, App
139	11.2	56.0	30	3	US-08-642-274D-211	Sequence 211, App
c 140	11.2	56.0	34	3	US-08-973-551-25	Sequence 25, Appl
c 141	11.2	56.0	34	4	US-09-565-808-6	Sequence 6, Appli
c 142	11.2	56.0	34	4	US-09-878-766A-7	Sequence 7, Appli
143	11.2	56.0	41	4	US-09-267-574-37	Sequence 37, Appl
144	11.2	56.0	42	1	US-08-621-897-8	Sequence 8, Appli
c 145	11.2	56.0	44	1	US-08-621-897-7	Sequence 7, Appli
146	11.2	56.0	45	1	US-08-206-176-19	Sequence 19, Appl
147	11.2	56.0	45	3	US-08-989-251-7	Sequence 7, Appli
148	11.2	56.0	45	3	US-09-340-250-7	Sequence 7, Appli
149	11.2	56.0	45	4	US-09-528-108-7	Sequence 7, Appli
c 150	11.2	56.0	47	4	US-09-671-317-694	Sequence 694, App

ALIGNMENTS

```
RESULT 1
US-09-359-757-34
; Sequence 34, Application US/09359757
; Patent No. 6080546
; GENERAL INFORMATION:
;  APPLICANT: Brett P. Monia
;  APPLICANT: William Gaarde
;  APPLICANT: Lex M. Cowsert
;  TITLE OF INVENTION: ANTISENSE MODULATION OF MEKK5 EXPRESSION
;  FILE REFERENCE: RTS-0078
;  CURRENT APPLICATION NUMBER: US/09/359,757
;  CURRENT FILING DATE: 1999-07-23
;  NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 34
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Antisense Oligonucleotide
US-09-359-757-34

  Query Match             69.0%;  Score 13.8;  DB 3;  Length 20;
  Best Local Similarity   88.2%;  Pred. No. 5.3e+02;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTG 17
              ||||||||||  | ||||
Db          2 AGTAACATCTGTCTTTG 18


RESULT 2
US-09-187-050-9
; Sequence 9, Application US/09187050B
; Patent No. 6043072
; GENERAL INFORMATION:
;  APPLICANT: Croteau, Rodney B
;  APPLICANT: Hefner, Jerry
;  TITLE OF INVENTION: Nucleic Acids Encoding Taxus Geranylgeranyl Diphosphate
;  TITLE OF INVENTION: Synthase, And Methods of Use
;  FILE REFERENCE: WSUR12423
;  CURRENT APPLICATION NUMBER: US/09/187,050B
;  CURRENT FILING DATE: 1998-11-05
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 43
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:PCR primer
;   FEATURE:
;   NAME/KEY: misc_difference
;   LOCATION: (1)..(43)
;   OTHER INFORMATION: PCR primer for synthesizing Tr295 truncation
;   OTHER INFORMATION: product
US-09-187-050-9

  Query Match             69.0%;  Score 13.8;  DB 3;  Length 43;
  Best Local Similarity   88.2%;  Pred. No. 5.6e+02;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 AACATCTATGTTTGGTT 20
              || |||||||||| ||
Db          2 AAGATCTATGTTTGATT 18


RESULT 3
US-08-559-303B-14
; Sequence 14, Application US/08559303B
; Patent No. 5824501
;  GENERAL INFORMATION:
;    APPLICANT:  NATHAN A. ELLIS, JAMES GERMAN, AND JOANNA
;    APPLICANT:  GRODEN
;    TITLE OF INVENTION:  METHODS FOR DIAGNOSIS AND TREATMENT
;    TITLE OF INVENTION:  OF BLOOM'S SYNDROME
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AMSTER, ROTHSTEIN & EBENSTEIN
;      STREET:  90 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  U.S.A.
;      ZIP:  10016
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH 1.44 Mb STORAGE DISKETTE
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/559,303B
;      FILING DATE:  NOVEMBER 15, 1995
```

```
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  ELIZABETH A. BOGOSIAN
;       REGISTRATION NUMBER:  39,911
;       REFERENCE/DOCKET NUMBER:  63475/65
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 697-5995
;       TELEFAX:  (212) 286-0854 or 286-0082
;       TELEX:  TWX 710-581-4766
;   INFORMATION FOR SEQ ID NO:  14:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  24
;       TYPE:  NUCLEIC ACID
;       STRANDEDNESS:  SINGLE
;       TOPOLOGY:  LINEAR
;     MOLECULE TYPE:
;       DESCRIPTION:  OTHER NUCLEIC ACID
;     HYPOTHETICAL:  YES
;     ANTI-SENSE:  NO
;     FEATURE:
;       NAME/KEY:
;       LOCATION:
;       IDENTIFICATION METHOD:
;       OTHER INFORMATION:
US-08-559-303B-14

  Query Match             66.0%;  Score 13.2;  DB 1;  Length 24;
  Best Local Similarity   83.3%;  Pred. No. 1e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGG 18
              |||| |||| ||| ||||
Db          6 AGTACCATCAATGATTGG 23


RESULT 4
US-09-175-828-14
; Sequence 14, Application US/09175828
; Patent No. 6221643
;  GENERAL INFORMATION:
;    APPLICANT:  NATHAN A. ELLIS, JAMES GERMAN, AND JOANNA
;    APPLICANT:  GRODEN
;    TITLE OF INVENTION:  METHODS FOR DIAGNOSIS AND TREATMENT
;    TITLE OF INVENTION:  OF BLOOM'S SYNDROME
;    NUMBER OF SEQUENCES:  78
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AMSTER, ROTHSTEIN & EBENSTEIN
;      STREET:  90 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  U.S.A.
;      ZIP:  10016
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 INCH 1.44 Mb STORAGE DISKETTE
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/175,828
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/559,303
;      FILING DATE:  NOVEMBER 15, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ELIZABETH A. BOGOSIAN
;      REGISTRATION NUMBER:  39,911
;      REFERENCE/DOCKET NUMBER:  63475/65
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 697-5995
;      TELEFAX:  (212) 286-0854 or 286-0082
;      TELEX:  TWX 710-581-4766
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:
;      DESCRIPTION:  OTHER NUCLEIC ACID
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
US-09-175-828-14

  Query Match             66.0%;  Score 13.2;  DB 3;  Length 24;
  Best Local Similarity   83.3%;  Pred. No. 1e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGG 18
              |||| |||| ||| ||||
Db          6 AGTACCATCAATGATTGG 23


RESULT 5
US-09-460-421-3
; Sequence 3, Application US/09460421
; Patent No. 6670167
; GENERAL INFORMATION:
;  APPLICANT:  Chen, Ping
;  APPLICANT:  Anderson, Mark
;  APPLICANT:  Deng, Ya-Li
;  APPLICANT:  Gaur, Smita
;  APPLICANT:  Kan, Chen Chen
;  APPLICANT:  Luo, Chun
;  APPLICANT:  Lundgren, Karen
;  APPLICANT:  Margosiak, Steve
;  APPLICANT:  Nguyen, Binh
;  APPLICANT:  O'Connor, Patrick
;  APPLICANT:  Register, James
;  APPLICANT:  Russell, Anna Tempczyk
;  APPLICANT:  Sarup, Jay
;  TITLE OF INVENTION:  Catalytic Domain of the Human Effector Cell cycle
;  TITLE OF INVENTION:  Checkpoint Protein Kinase, Chk1, Materials and
;  TITLE OF INVENTION:  Methods for Identification of Inhibitors Thereof
;  FILE REFERENCE:  0125-0032
;  CURRENT APPLICATION NUMBER: US/09/460,421
;  CURRENT FILING DATE:  1999-12-14
;  NUMBER OF SEQ ID NOS:  24
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 34
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: PCR primer
US-09-460-421-3

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 34;
  Best Local Similarity   87.5%;  Pred. No. 1.7e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTT 16
              |||| ||||||| |||
Db          7 AGTACCATCTATCTTT 22


RESULT 6
US-09-422-978-1887/c
; Sequence 1887, Application US/09422978
; Patent No. 6537751
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT: Blumenfeld, Marta
```

```
;  APPLICANT: Chumakov, Ilya
;  TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
;  FILE REFERENCE: GENSET.020CP1
;  CURRENT APPLICATION NUMBER: US/09/422,978
;  CURRENT FILING DATE: 1999-10-20
;  EARLIER APPLICATION NUMBER: US 09/298,850
;  EARLIER FILING DATE: 1999-04-21
;  EARLIER APPLICATION NUMBER: US 60/109,732
;  EARLIER FILING DATE: 1998-11-23
;  EARLIER APPLICATION NUMBER: US 60/082,614
;  EARLIER FILING DATE: 1998-04-21
;  NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 1887
;    LENGTH: 47
;    TYPE: DNA
;    ORGANISM: Homo Sapiens
;    FEATURE:
;    NAME/KEY: allele
;    LOCATION: 24
;    OTHER INFORMATION: 99-7129-335 : polymorphic base A or C
US-09-422-978-1887

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 47;
  Best Local Similarity   87.5%;  Pred. No. 1.7e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGGTT 20
              |||| |||||||| ||
Db         43 ACATTTATGTTTGTTT 28


RESULT 7
US-09-315-793-53/c
; Sequence 53, Application US/09315793
; Patent No. 6221597
; GENERAL INFORMATION:
;  APPLICANT: Roberts, Christopher J.
;  TITLE OF INVENTION: ESSENTIAL GENES OF YEAST AS TARGETS FOR ANTIFUNGAL
;  TITLE OF INVENTION: AGENTS, HERBICIDES, INSECTICIDES AND ANTI-PROLIFERATION
;  TITLE OF INVENTION: DRUGS
;  FILE REFERENCE: 9301-048
;  CURRENT APPLICATION NUMBER: US/09/315,793
;  CURRENT FILING DATE: 1999-05-21
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 53
;    LENGTH: 74
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence: primer
US-09-315-793-53

  Query Match             64.0%;  Score 12.8;  DB 3;  Length 74;
  Best Local Similarity   87.5%;  Pred. No. 1.8e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGGTT 20
              ||||| || |||||||
Db         23 ACATCCATCTTTGGTT 8


RESULT 8
US-08-868-699A-5/c
; Sequence 5, Application US/08868699A
; Patent No. 6204019
;  GENERAL INFORMATION:
;    APPLICANT:  O'Dwyer, Karen
;    APPLICANT:  Perry, Caroline
;    APPLICANT:  Warren, Richard L.
;    TITLE OF INVENTION:  No. 6204019el Compounds
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dechert, Price & Rhoads
;      STREET:  4000 Bell Atlantic Tower, 1717 Arch Stre
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103-2793
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/868,699A
;      FILING DATE:  04-JUN-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Falk, Stephen T
;      REGISTRATION NUMBER:  36,795
;      REFERENCE/DOCKET NUMBER:  GM10012
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-994-2488
;      TELEFAX:  215-994-2222
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-868-699A-5

  Query Match             63.0%;  Score 12.6;  DB 3;  Length 24;
  Best Local Similarity   78.9%;  Pred. No. 2e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGTT 20
              |||||||||| ||  |||
Db         23 GTAACATCTAGTTTATGTT 5


RESULT 9
US-09-757-014-5/c
; Sequence 5, Application US/09757014
; Patent No. 6348342
;   GENERAL INFORMATION:
;        APPLICANT: O'Dwyer, Karen
;                   Perry, Caroline
;                   Warren, Richard L.
;        TITLE OF INVENTION: No. 6348342el Compounds
;        NUMBER OF SEQUENCES: 6
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Dechert, Price & Rhoads
;             STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
;             CITY: Philadelphia
;             STATE: PA
;             COUNTRY: USA
;             ZIP: 19103-2793
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: Windows
;             SOFTWARE: FastSEQ for Windows Version 2.0b
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/757,014
;             FILING DATE: 09-Jan-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/868,699
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
```

```
;            NAME: Falk, Stephen T
;            REGISTRATION NUMBER: 36,795
;            REFERENCE/DOCKET NUMBER: GM10012
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 215-994-2488
;            TELEFAX: 215-994-2222
;            TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 5:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 24 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-757-014-5

  Query Match           63.0%;  Score 12.6;  DB 4;  Length 24;
  Best Local Similarity  78.9%;  Pred. No. 2e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGTT 20
              ||||||||||  ||  |||
Db         23 GTAACATCTAGTTTATGTT 5


RESULT 10
US-09-009-913-223/c
; Sequence 223, Application US/09009913
; Patent No. 6087485
;  GENERAL INFORMATION:
;    APPLICANT:  AxyS Pharmaceuticals, Inc.
;    TITLE OF INVENTION:  Asthma Related Genes
;    NUMBER OF SEQUENCES:  339
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Bozicevic & Reed, LLP
;      STREET:  285 Hamilton Ave, Suite 200
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/009,913
;      FILING DATE:  21-JAN-1998
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sherwood, Pamela J
;      REGISTRATION NUMBER:  36,677
;      REFERENCE/DOCKET NUMBER:  SEQ-4P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-327-3231
;      TELEFAX:  650-327-3231
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  223:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-009-913-223

  Query Match           63.0%;  Score 12.6;  DB 3;  Length 27;
  Best Local Similarity  78.9%;  Pred. No. 2e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              |||||||||| |  ||||
Db         27 AGTAACATCTCAGCCTGGT 9


RESULT 11
US-07-749-446-3
; Sequence 3, Application US/07749446
; Patent No. 5593857
;  GENERAL INFORMATION:
;    APPLICANT:  Higaki, Jeffrey N.
;    APPLICANT:  Tischer, Edmund G.
;    APPLICANT:  Cordell, Barbara
;    APPLICANT:  Thompson, Stewart A.
;    TITLE OF INVENTION:  PRODUCTION OF HOMOGENEOUS CILIARY
;    TITLE OF INVENTION:  NEUROTROPHIC FACTOR
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  California Biotechnology Inc.
;      STREET:  2450 Bayshore Parkway
;      CITY:  Mountain View
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94043
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/749,446
;      FILING DATE:  19911008
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shearer, Peter R.
;      REGISTRATION NUMBER:  28,117
;      REFERENCE/DOCKET NUMBER:  PC43:US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-962-5860
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  34 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-07-749-446-3

  Query Match           63.0%;  Score 12.6;  DB 1;  Length 34;
  Best Local Similarity  78.9%;  Pred. No. 2.1e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGTT 20
              ||| | || ||||||  |||
Db         15 GTACCTTCCATGTTTTGTT 33


RESULT 12
US-08-584-760A-50
; Sequence 50, Application US/08584760A
; Patent No. 6290953
;  GENERAL INFORMATION:
;    APPLICANT:  Ballance, David J
;    APPLICANT:  Courtney, Michael G
;    APPLICANT:  Finnis, Christopher J A
;    APPLICANT:  Sleep, Darrell
;    TITLE OF INVENTION:  Medicine
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Centeon L.L.C.
;      STREET:  1020 First Avenue
;      CITY:  King of Prussia
;      STATE:  Pennsylvania
;      COUNTRY:  USA
```

```
;       ZIP:  19406-1310
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/584,760A
;       FILING DATE:
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/211,860
;       FILING DATE:  15-APR-1994
;       APPLICATION NUMBER:  GB 9121815.6
;       FILING DATE:  14-OCT-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Naomi Biswas
;       REGISTRATION NUMBER:  38,384
;       REFERENCE/DOCKET NUMBER:  92H853-1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  610/878-4294
;       TELEFAX:  610/878/4221
;   INFORMATION FOR SEQ ID NO:  50:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  35 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  1..35
;       OTHER INFORMATION:  /function= "oligonucleotide 50"
US-08-584-760A-50

  Query Match             63.0%;  Score 12.6;  DB 3;  Length 35;
  Best Local Similarity   78.9%;  Pred. No. 2.1e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              | ||||||||| || ||| |
Db          5 AATAACATCTTTGCTTGTT 23


RESULT 13
US-08-584-760A-49/c
; Sequence 49, Application US/08584760A
; Patent No. 6290953
;  GENERAL INFORMATION:
;     APPLICANT:  Ballance, David J
;     APPLICANT:  Courtney, Michael G
;     APPLICANT:  Finnis, Christopher J A
;     APPLICANT:  Sleep, Darrell
;     TITLE OF INVENTION:  Medicine
;     NUMBER OF SEQUENCES:  76
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Centeon L.L.C.
;       STREET:  1020 First Avenue
;       CITY:  King of Prussia
;       STATE:  Pennsylvania
;       COUNTRY:  USA
;       ZIP:  19406-1310
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/584,760A
;       FILING DATE:
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/211,860
;       FILING DATE:  15-APR-1994
;       APPLICATION NUMBER:  GB 9121815.6
;       FILING DATE:  14-OCT-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Naomi Biswas
;       REGISTRATION NUMBER:  38,384
;       REFERENCE/DOCKET NUMBER:  92H853-1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  610/878-4294
;       TELEFAX:  610/878/4221
;   INFORMATION FOR SEQ ID NO:  49:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  36 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  1..36
;       OTHER INFORMATION:  /function= "oligonucleotide 49"
US-08-584-760A-49

  Query Match             63.0%;  Score 12.6;  DB 3;  Length 36;
  Best Local Similarity   78.9%;  Pred. No. 2.1e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              | ||||||||| || ||| |
Db         35 AATAACATCTTTGCTTGTT 17


RESULT 14
US-09-086-726-11/c
; Sequence 11, Application US/09086726
; Patent No. 6346378
; GENERAL INFORMATION:
;  APPLICANT: Stanley, Christopher John
;  APPLICANT:  Orum, Henrik
;  APPLICANT:  Jorgensen, Mikkel
;  TITLE OF INVENTION: Nucleic Acid Analogs With A Chelating Functionality
;  FILE REFERENCE: 108382-08046
;  CURRENT APPLICATION NUMBER: US/09/086,726
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 08/653,607
;  PRIOR FILING DATE: 1996-05-24
;  PRIOR APPLICATION NUMBER: PCT/EP94/03859
;  PRIOR FILING DATE: 1994-11-22
;  PRIOR APPLICATION NUMBER: UK/9324243.6
;  PRIOR FILING DATE: 1993-11-25
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
;   LENGTH: 37
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: The sequence of an intramolecular stem structure of an
;   OTHER INFORMATION: oligonucleotide
US-09-086-726-11

  Query Match             63.0%;  Score 12.6;  DB 4;  Length 37;
  Best Local Similarity   78.9%;  Pred. No. 2.1e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGTT 20
              || ||| |||| ||| |||
Db         32 GTCACAACTATTTTTAGTT 14
```

```
RESULT 15
US-08-233-130A-3/c
; Sequence 3, Application US/08233130A
; Patent No. 5587300
;  GENERAL INFORMATION:
;    APPLICANT:  Malter, James S.
;    TITLE OF INVENTION:  Method to Increase Regulatory Molecule
;    TITLE OF INVENTION:  Production
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Mueting, Raasch, Gebhardt & Schwappach, P.A.
;      STREET:  203 Textile Buildling, 119 No. 5587300th Fourth Street
;      CITY:  Minneapolis
;      STATE:  MN
;      COUNTRY:  USA
;      ZIP:  55401
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/233,130A
;      FILING DATE:  26-APR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mueting, Ann M.
;      REGISTRATION NUMBER:  33,977
;      REFERENCE/DOCKET NUMBER:  119.00010101
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  612-305-1220
;      TELEFAX:  612-305-1228
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  34 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-233-130A-3

  Query Match             61.0%;  Score 12.2;  DB 1;  Length 34;
  Best Local Similarity   82.4%;  Pred. No. 3.2e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTG 17
              ||||| || ||||| ||
Db         23 AGTAATATGTATGTATG 7


RESULT 16
US-09-270-140A-52
; Sequence 52, Application US/09270140A
; Patent No. 6361941
; GENERAL INFORMATION:
;  APPLICANT: Todd, Alison
;  APPLICANT:  Fuery, Caroline
;  APPLICANT:  Cairns, Murray
;  TITLE OF INVENTION: Catalytic Nucleic Acid base Diagnostic Methods
;  FILE REFERENCE: J&J1799
;  CURRENT APPLICATION NUMBER: US/09/270,140A
;  CURRENT FILING DATE:  1999-03-16
;  PRIOR APPLICATION NUMBER: 60/079,651
;  PRIOR FILING DATE: 1998-03-27
;  NUMBER OF SEQ ID NOS: 96
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 52
;   LENGTH: 22
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:wildtype RNA
;   OTHER INFORMATION: for Codon 508 - wildtype
US-09-270-140A-52

  Query Match             60.0%;  Score 12;  DB 4;  Length 22;
  Best Local Similarity   35.0%;  Pred. No. 3.9e+03;
  Matches    7;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGTT 20
              | :| ||:|: :| ::| ::
Db          3 AAUAUCAUCUUUGGUUGUUU 22


RESULT 17
US-08-629-001A-60
; Sequence 60, Application US/08629001A
; Patent No. 5858661
;  GENERAL INFORMATION:
;    APPLICANT:  Shiloh, Yosef
;    TITLE OF INVENTION:  ATAXIA-TELANGIECTASIA GENE AND ITS
;    TITLE OF INVENTION:  GENOMIC ORGANIZATION
;    NUMBER OF SEQUENCES:  139
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kohn & Associates
;      STREET:  30500 No. 5858661thwestern Hwy.
;      CITY:  Farmington Hills
;      STATE:  Michigan
;      COUNTRY:  US
;      ZIP:  48334
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/629,001A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kohn, Kenneth I.
;      REGISTRATION NUMBER:  30,955
;      REFERENCE/DOCKET NUMBER:  2290.00032
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (810) 539-5050
;      TELEFAX:  (810) 539-5055
;  INFORMATION FOR SEQ ID NO:  60:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-629-001A-60

  Query Match             60.0%;  Score 12;  DB 2;  Length 30;
  Best Local Similarity   75.0%;  Pred. No. 4e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGTT 20
              |||||||| ||| |   |||
Db         10 AGTAACATGTATTTGCTGTT 29


RESULT 18
US-08-642-274D-139
; Sequence 139, Application US/08642274D
; Patent No. 6200749
; GENERAL INFORMATION:
;  APPLICANT:  Shiloh, Yosef
;  TITLE OF INVENTION:  MUTATED FORMS OF THE ATAXIA-TELANGIECTASIA GENE AND METHOD TO
;  TITLE OF INVENTION:  SCREEN FOR A PARTIAL A-T PHENOTYPE
;  FILE REFERENCE:  229000033
;  CURRENT APPLICATION NUMBER: US/08/642,274D
;  CURRENT FILING DATE:  1996-05-03
;  NUMBER OF SEQ ID NOS:  220
```

```
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 139
;    LENGTH: 30
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence:intronic
;    OTHER INFORMATION: sequence
US-08-642-274D-139

  Query Match             60.0%;  Score 12;  DB 3;  Length 30;
  Best Local Similarity   75.0%;  Pred. No. 4e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGTT 20
              ||||||||  ||| |   |||
Db         10 AGTAACATGTATTTGCTGTT 29


RESULT 19
US-08-233-009-51
; Sequence 51, Application US/08233009
; Patent No. 5646156
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobson, Marlene A
;    APPLICANT:  Johnson, Robert G
;    APPLICANT:  Salvatore, Christopher A
;    TITLE OF INVENTION:  INHIBITION OF EOSINOPHIL
;    TITLE OF INVENTION:  ACTIVATION THROUGH A3 ADENOSINE RECEPTOR ANTAGONISM
;    NUMBER OF SEQUENCES:  56
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  P.O.Box 2000
;      CITY:  Rahway
;      STATE:  New Jersey
;      COUNTRY:  United States
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/233,009
;      FILING DATE:  25-APR-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bencen, Gerard H
;      REGISTRATION NUMBER:  35,746
;      REFERENCE/DOCKET NUMBER:  19219
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (908) 594-3901
;      TELEFAX:  (908)594-4720
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  YES
US-08-233-009-51

  Query Match             60.0%;  Score 12;  DB 1;  Length 45;
  Best Local Similarity   75.0%;  Pred. No. 4.1e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGTT 20
              | | ||  |||||||||| |
Db          8 ACTGACCCCTATGTTTGGCT 27


RESULT 20
US-09-422-978-2712
; Sequence 2712, Application US/09422978
; Patent No. 6537751
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT: Blumenfeld, Marta
;  APPLICANT: Chumakov, Ilya
;  TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
;  FILE REFERENCE: GENSET.020CP1
;  CURRENT APPLICATION NUMBER: US/09/422,978
;  CURRENT FILING DATE: 1999-10-20
;  EARLIER APPLICATION NUMBER: US 09/298,850
;  EARLIER FILING DATE: 1999-04-21
;  EARLIER APPLICATION NUMBER: US 60/109,732
;  EARLIER FILING DATE: 1998-11-23
;  EARLIER APPLICATION NUMBER: US 60/082,614
;  EARLIER FILING DATE: 1998-04-21
;  NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 2712
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: 99-15423-223 : polymorphic base G or A
US-09-422-978-2712

  Query Match             60.0%;  Score 12;  DB 4;  Length 47;
  Best Local Similarity   85.7%;  Pred. No. 4.1e+03;
  Matches   12;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTT 15
              :||||||||||  ||
Db         24 RTAACATCTATATT 37


RESULT 21
US-09-443-199C-847/c
; Sequence 847, Application US/09443199C
; Patent No. 6670464
; GENERAL INFORMATION:
;  APPLICANT: Shimkets, Richard A.
;  APPLICANT:  Leach, Martin
;  TITLE OF INVENTION: Nucleic Acids Containing Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms and Methods of Use Thereof
;  FILE REFERENCE: 15966-534A
;  CURRENT APPLICATION NUMBER: US/09/443,199C
;  CURRENT FILING DATE:  1999-11-16
;  PRIOR APPLICATION NUMBER: 60/109,024
;  PRIOR FILING DATE: 1998-11-17
;  NUMBER OF SEQ ID NOS: 1272
;  SOFTWARE: CuraGen Patent Formatter Version 0.9
; SEQ ID NO 847
;   LENGTH: 51
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (26)...(0)
;   OTHER INFORMATION: 1 of 2 allelic variants (848 is other entry)
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Accession number cg43949585
US-09-443-199C-847

  Query Match             60.0%;  Score 12;  DB 4;  Length 51;
  Best Local Similarity   75.0%;  Pred. No. 4.2e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGTT 20
              | || |||||  | ||| |||
```

```
Db         30 ATTATCATCTGTATTTAGTT 11


RESULT 22
US-09-443-199C-848/c
; Sequence 848, Application US/09443199C
; Patent No. 6670464
; GENERAL INFORMATION:
;  APPLICANT: Shimkets, Richard A.
;  APPLICANT:  Leach, Martin
;  TITLE OF INVENTION: Nucleic Acids Containing Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms and Methods of Use Thereof
;  FILE REFERENCE: 15966-534A
;  CURRENT APPLICATION NUMBER: US/09/443,199C
;  CURRENT FILING DATE:  1999-11-16
;  PRIOR APPLICATION NUMBER: 60/109,024
;  PRIOR FILING DATE: 1998-11-17
;  NUMBER OF SEQ ID NOS: 1272
;  SOFTWARE: CuraGen Patent Formatter Version 0.9
; SEQ ID NO 848
;   LENGTH: 51
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (26)...(0)
;   OTHER INFORMATION: 2 of 2 allelic variants (847 is other entry)
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Accession number cg43949585
US-09-443-199C-848    .

  Query Match             60.0%;  Score 12;  DB 4;  Length 51;
  Best Local Similarity   75.0%;  Pred. No. 4.2e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGTT 20
              | || ||||| | ||| |||
Db         30 ATTACCATCTGTATTTAGTT 11


RESULT 23
US-08-864-473-61
; Sequence 61, Application US/08864473
; Patent No. 6027889
; GENERAL INFORMATION:
;  APPLICANT: Barany, Francis
;  APPLICANT: Lubin, Matthew
;  TITLE OF INVENTION: DETECTION OF NUCLEIC ACID SEQUENCE DIFFERENCES USING
;  TITLE OF INVENTION: COUPLED LIGASE DETECTION AND POLYMERASE CHAIN REACTIONS
;  FILE REFERENCE: 19603/441
;  CURRENT APPLICATION NUMBER: US/08/864,473
;  CURRENT FILING DATE: 1997-05-28
;  EARLIER APPLICATION NUMBER: 60/018,532
;  EARLIER FILING DATE: 1996-05-29
;  NUMBER OF SEQ ID NOS: 76
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61
;   LENGTH: 57
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Primer
;   OTHER INFORMATION: Sequence
US-08-864-473-61

  Query Match             60.0%;  Score 12;  DB 3;  Length 57;
  Best Local Similarity   100.0%;  Pred. No. 4.2e+03;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TAACATCTATGT 14
              ||||||||||||
Db         29 TAACATCTATGT 40


RESULT 24
US-09-440-523-61
; Sequence 61, Application US/09440523
; Patent No. 6268148
; GENERAL INFORMATION:
;  APPLICANT:  Barany, Francis
;  APPLICANT:  Lubin, Matthew
;  TITLE OF INVENTION:  DETECTION OF NUCLEIC ACID SEQUENCE DIFFERENCES USING
;  TITLE OF INVENTION:  COUPLED LIGASE DETECTION AND POLYMERASE CHAIN REACTIONS
;  FILE REFERENCE:  19603/441
;  CURRENT APPLICATION NUMBER: US/09/440,523
;  CURRENT FILING DATE:  1999-11-15
;  PRIOR APPLICATION NUMBER:  08/864,473
;  PRIOR FILING DATE:  1997-05-28
;  NUMBER OF SEQ ID NOS:  76
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 61
;   LENGTH: 57
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Primer
;   OTHER INFORMATION: Sequence
US-09-440-523-61

  Query Match             60.0%;  Score 12;  DB 3;  Length 57;
  Best Local Similarity   100.0%;  Pred. No. 4.2e+03;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TAACATCTATGT 14
              ||||||||||||
Db         29 TAACATCTATGT 40


RESULT 25
US-08-956-171E-4982/c
; Sequence 4982, Application US/08956171E
; Patent No. 6593114
;   GENERAL INFORMATION:
;        APPLICANT: Charles Kunsch
;                   Gil H. Choi
;                   Patrick S. Dillon
;                   Craig A. Rosen
;                   Steven C. Barash
;                   Michael R. Fannon
;        TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;        NUMBER OF SEQUENCES: 5256
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/956,171E
;             FILING DATE: 20-Oct-1997
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/009,861
;             FILING DATE: January 5, 1996
;             APPLICATION NUMBER: 08/781,986
;             FILING DATE: January 3, 1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Mark J. Hyman
;             REGISTRATION NUMBER: 46,789
```

```
;           REFERENCE/DOCKET NUMBER: PB248P1
;       TELECOMMUNICATION INFORMATION:
;           TELEPHONE: (240) 314-1224
;           TELEFAX: (301) 309-8439
;   INFORMATION FOR SEQ ID NO: 4982:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 58 base pairs
;           TYPE: nucleic acid
;           STRANDEDNESS: double
;           TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 4982:
US-08-956-171E-4982

  Query Match             60.0%;  Score 12;  DB 4;  Length 58;
  Best Local Similarity   75.0%;  Pred. No. 4.2e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGTT 20
              ||  | ||||||| |||| |
Db         55 AGATAAATCTATGATTGGAT 36


RESULT 26
US-08-303-275-198
; Sequence 198, Application US/08303275
; Patent No. 5766598
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    APPLICANT:  Tartaglia, James
;    APPLICANT:  Cox, William I.
;    TITLE OF INVENTION:  IMMUNODEFICIENCY VIRUS RECOMBINANT
;    TITLE OF INVENTION:  POXVIRUS VACCINE
;    NUMBER OF SEQUENCES:  205
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o William S. Frommer
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/303,275
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/897,382
;      FILING DATE:  11-JUN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2420
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  198:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  63 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-303-275-198

  Query Match             60.0%;  Score 12;  DB 1;  Length 63;
  Best Local Similarity   75.0%;  Pred. No. 4.2e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGTT 20
              || || | |||||| || ||
Db         42 AGAAAAAGCTATGTATGCTT 61


RESULT 27
US-08-303-275-199/c
; Sequence 199, Application US/08303275
; Patent No. 5766598
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    APPLICANT:  Tartaglia, James
;    APPLICANT:  Cox, William I.
;    TITLE OF INVENTION:  IMMUNODEFICIENCY VIRUS RECOMBINANT
;    TITLE OF INVENTION:  POXVIRUS VACCINE
;    NUMBER OF SEQUENCES:  205
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o William S. Frommer
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/303,275
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/897,382
;      FILING DATE:  11-JUN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2420
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  199:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  63 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-303-275-199

  Query Match             60.0%;  Score 12;  DB 1;  Length 63;
  Best Local Similarity   75.0%;  Pred. No. 4.2e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGTT 20
              || || | |||||| || ||
Db         22 AGAAAAAGCTATGTATGCTT 3


RESULT 28
US-09-479-005A-399
; Sequence 399, Application US/09479005A
; Patent No. 6656731
; GENERAL INFORMATION:
;  APPLICANT: Ribozyme Pharmaceuticals, Inc.
;  TITLE OF INVENTION: Nucleic Acid Catalysts with Endonuclease Activity
;  FILE REFERENCE: MBHB00-884-C
;  CURRENT APPLICATION NUMBER: US/09/479,005A
;  CURRENT FILING DATE:  2000-01-07
;  PRIOR APPLICATION NUMBER: US 09/444,209
;  PRIOR FILING DATE: 1999-11-19
;  PRIOR APPLICATION NUMBER: US 09/159,274
;  PRIOR FILING DATE: 1998-09-22
```

```
;  PRIOR APPLICATION NUMBER: US 60/059,473
;  PRIOR FILING DATE: 1997-09-22
;  NUMBER OF SEQ ID NOS: 1208
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 399
;   LENGTH: 16
;   TYPE: RNA
;   ORGANISM: Homo sapiens
US-09-479-005A-399

  Query Match             59.0%;  Score 11.8;  DB 4;  Length 16;
  Best Local Similarity   46.7%;  Pred. No. 4.8e+03;
  Matches    7;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTT 15
              ||:||| :|: :|::
Db          1 AGUAACUUCUGUGUU 15


RESULT 29
US-08-261-822A-57
; Sequence 57, Application US/08261822A
; Patent No. 5650553
;  GENERAL INFORMATION:
;    APPLICANT:  Ecker, Joseph R. et al.
;    TITLE OF INVENTION:  Plant Genes for Sensitivity to Ethylene
;    TITLE OF INVENTION:  and Pathogens
;    NUMBER OF SEQUENCES:  82
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock, Washburn, Kurtz, Mackiewicz & No. 5650553ris
;      STREET:  One Liberty Place, 46th floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/261,822A
;      FILING DATE:  17-JUN-1994
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Beardell, Lori Y.
;      REGISTRATION NUMBER:  34,293
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  57:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  18 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  YES
US-08-261-822A-57

  Query Match             59.0%;  Score 11.8;  DB 1;  Length 18;
  Best Local Similarity   86.7%;  Pred. No. 4.9e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          6 CATCTATGTTTGGTT 20
              | ||||| |||||||
Db          2 CTTCTATATTTGGTT 16


RESULT 30
PCT-US95-07744A-57
; Sequence 57, Application PC/TUS9507744A
;  GENERAL INFORMATION:
;    APPLICANT:  Trustees of The University of Pennsylvania
;    TITLE OF INVENTION:  Plant Genes for Sensitivity to Ethylene
;    TITLE OF INVENTION:  and Pathogens
;    NUMBER OF SEQUENCES:  82
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock, Washburn, Kurtz, Mackiewicz & Norris
;      STREET:  One Liberty Place, 46th floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/07744A
;      FILING DATE:  15-JUNE-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/261,822
;      FILING DATE:  June 17, 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Beardell, Lori Y.
;      REGISTRATION NUMBER:  34,293
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  57:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  18 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  YES
PCT-US95-07744A-57

  Query Match             59.0%;  Score 11.8;  DB 5;  Length 18;
  Best Local Similarity   86.7%;  Pred. No. 4.9e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          6 CATCTATGTTTGGTT 20
              | ||||| |||||||
Db          2 CTTCTATATTTGGTT 16


RESULT 31
US-09-287-796-149/c
; Sequence 149, Application US/09287796A
; Patent No. 6133246
; GENERAL INFORMATION:
;  APPLICANT: McKay, Robert A.
;  APPLICANT: Dean, Nicholas M.
;  APPLICANT: Monia, Brett
;  APPLICANT: Nero, Pam
;  APPLICANT: Gaarde, William A.
;  TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE COMPOSITIONS AND METHODS
;  TITLE OF INVENTION: FOR THE MODULATION OF JNK PROTEINS
;  FILE REFERENCE: ISPH-0350
;  CURRENT APPLICATION NUMBER: US/09/287,796A
;  CURRENT FILING DATE: 1999-04-07
;  EARLIER APPLICATION NUMBER: 09/130,616
;  EARLIER FILING DATE: 1998-08-07
;  EARLIER APPLICATION NUMBER: 08/910,629
;  EARLIER FILING DATE: 1997-08-03
;  NUMBER OF SEQ ID NOS: 165
; SEQ ID NO 149
;   LENGTH: 20
;   TYPE: DNA
```

```
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic Sequence
US-09-287-796-149

  Query Match             59.0%;  Score 11.8;  DB 3;  Length 20;
  Best Local Similarity   86.7%;  Pred. No. 4.9e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGGT 19
              ||||| | |||||||
Db         19 ACATCAACGTTTGGT 5


RESULT 32
US-09-130-616-149/c
; Sequence 149, Application US/09130616C
; Patent No. 6221850
; GENERAL INFORMATION:
;  APPLICANT: McKay, Robert A.
;  APPLICANT: Dean, Nicholas M.
;  APPLICANT: Monia, Brett
;  APPLICANT: Nero, Pam
;  APPLICANT: Gaarde, William A.
;  TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE COMPOSITIONS AND METHODS
;  TITLE OF INVENTION: FOR THE MODULATION OF JNK PROTEINS
;  FILE REFERENCE: ISPH-0318
;  CURRENT APPLICATION NUMBER: US/09/130,616C
;  CURRENT FILING DATE: 1998-08-07
;  EARLIER APPLICATION NUMBER: 08/910,629
;  EARLIER FILING DATE: 1997-08-03
;  NUMBER OF SEQ ID NOS: 178
; SEQ ID NO 149
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic sequence
US-09-130-616-149

  Query Match             59.0%;  Score 11.8;  DB 3;  Length 20;
  Best Local Similarity   86.7%;  Pred. No. 4.9e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGGT 19
              ||||| | |||||||
Db         19 ACATCAACGTTTGGT 5


RESULT 33
US-08-445-746-10
; Sequence 10, Application US/08445746
; Patent No. 5709865
;  GENERAL INFORMATION:
;    APPLICANT:  Jan van den Hurk and Peter Tijssen
;    TITLE OF INVENTION:  Bovine Viral Diarrhea Virus Group II
;    TITLE OF INVENTION:  gp53 Compositions and Methods
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  350 Cambridge Avenue, Suite 250
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/445,746
;      FILING DATE:  22-MAY-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/337,618
;      FILING DATE:  10-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sholtz, Charles K.
;      REGISTRATION NUMBER:  38,615
;      REFERENCE/DOCKET NUMBER:  1242-0001.30
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Primer 1A
US-08-445-746-10

  Query Match             59.0%;  Score 11.8;  DB 1;  Length 29;
  Best Local Similarity   86.7%;  Pred. No. 5e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTT 16
              | || ||||||||||
Db          1 GGAAGATCTATGTTT 15


RESULT 34
US-09-008-722-10
; Sequence 10, Application US/09008722
; Patent No. 6015795
;  GENERAL INFORMATION:
;    APPLICANT:  Jan van den Hurk and Peter Tijssen
;    TITLE OF INVENTION:  Bovine Viral Diarrhea Virus Group II
;    TITLE OF INVENTION:  gp53 Compositions and Methods
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  350 Cambridge Avenue, Suite 250
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/008,722
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/445,746
;      FILING DATE:  22-MAY-1995
;      APPLICATION NUMBER:  US 08/337,618
;      FILING DATE:  10-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sholtz, Charles K.
;      REGISTRATION NUMBER:  38,615
;      REFERENCE/DOCKET NUMBER:  1242-0001.30
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
```

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Primer 1A
US-09-008-722-10

  Query Match             59.0%;  Score 11.8;  DB 3;  Length 29;
  Best Local Similarity   86.7%;  Pred. No. 5e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTT 16
              | || ||||||||||
Db          1 GGAAGATCTATGTTT 15


RESULT 35
US-08-320-373-1
; Sequence 1, Application US/08320373
; Patent No. 5559025
;  GENERAL INFORMATION:
;    APPLICANT:  Ahorn, Horst
;    APPLICANT:  Maurer-Fogy, Ingrid
;    APPLICANT:  Sommergruber, Wolfgang
;    APPLICANT:  Zophel, Andreas
;    APPLICANT:  Blaas, Dieter
;    APPLICANT:  Kuchler, Ernst
;    APPLICANT:  Liebig, Hans-Dieter
;    APPLICANT:  Skern, Timothy
;    TITLE OF INVENTION:  Expression of Mature Proteinase 2A, the
;    TITLE OF INVENTION:  Partial Purification Thereof and Preparation of Substrates
;    TITLE OF INVENTION:  Having an Inhibitory Effect
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
;      STREET:  1225 Connecticut Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/320,373
;      FILING DATE:  11-OCT-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/971,619
;      FILING DATE:  06-NOV-1992
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 466-0800
;      TELEFAX:  (202) 833-8716
;      TELEX:  248636 SSK
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-320-373-1

  Query Match             59.0%;  Score 11.8;  DB 1;  Length 32;
  Best Local Similarity   86.7%;  Pred. No. 5e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTT 15
              ||| |||| ||||||
Db         11 AGTGACATGTATGTT 25


RESULT 36
US-08-320-373-2/c
; Sequence 2, Application US/08320373
; Patent No. 5559025
;  GENERAL INFORMATION:
;    APPLICANT:  Ahorn, Horst
;    APPLICANT:  Maurer-Fogy, Ingrid
;    APPLICANT:  Sommergruber, Wolfgang
;    APPLICANT:  Zophel, Andreas
;    APPLICANT:  Blaas, Dieter
;    APPLICANT:  Kuchler, Ernst
;    APPLICANT:  Liebig, Hans-Dieter
;    APPLICANT:  Skern, Timothy
;    TITLE OF INVENTION:  Expression of Mature Proteinase 2A, the
;    TITLE OF INVENTION:  Partial Purification Thereof and Preparation of Substrates
;    TITLE OF INVENTION:  Having an Inhibitory Effect
;    NUMBER OF SEQUENCES:  91
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
;      STREET:  1225 Connecticut Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/320,373
;      FILING DATE:  11-OCT-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/971,619
;      FILING DATE:  06-NOV-1992
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 466-0800
;      TELEFAX:  (202) 833-8716
;      TELEX:  248636 SSK
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  33 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-320-373-2

  Query Match             59.0%;  Score 11.8;  DB 1;  Length 33;
  Best Local Similarity   86.7%;  Pred. No. 5e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTT 15
              ||| |||| ||||||
Db         27 AGTGACATGTATGTT 13


RESULT 37
US-08-956-171E-2061/c
; Sequence 2061, Application US/08956171E
; Patent No. 6593114
;   GENERAL INFORMATION:
;        APPLICANT: Charles Kunsch
;                   Gil H. Choi
;                   Patrick S. Dillon
;                   Craig A. Rosen
;                   Steven C. Barash
;                   Michael R. Fannon
;        TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;        NUMBER OF SEQUENCES: 5256
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
```

```
;            STREET: 9410 Key West Avenue
;            CITY: Rockville
;            STATE: Maryland
;            COUNTRY: USA
;            ZIP: 20850
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;            COMPUTER: HP Vectra 486/33
;            OPERATING SYSTEM: MSDOS version 6.2
;            SOFTWARE: ASCII Text
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/956,171E
;            FILING DATE: 20-Oct-1997
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 60/009,861
;            FILING DATE: January 5, 1996
;            APPLICATION NUMBER: 08/781,986
;            FILING DATE: January 3, 1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Mark J. Hyman
;            REGISTRATION NUMBER: 46,789
;            REFERENCE/DOCKET NUMBER: PB248P1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (240) 314-1224
;            TELEFAX: (301) 309-8439
;   INFORMATION FOR SEQ ID NO: 2061:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 50 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2061:
US-08-956-171E-2061

  Query Match             59.0%;  Score 11.8;  DB 4;  Length 50;
  Best Local Similarity   86.7%;  Pred. No. 5.2e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 AACATCTATGTTTGG 18
              || |||| |||||||
Db         35 AATATCTTTGTTTGG 21


RESULT 38
US-08-956-171E-5031
; Sequence 5031, Application US/08956171E
; Patent No. 6593114
;   GENERAL INFORMATION:
;       APPLICANT: Charles Kunsch
;                  Gil H. Choi
;                  Patrick S. Dillon
;                  Craig A. Rosen
;                  Steven C. Barash
;                  Michael R. Fannon
;       TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;       NUMBER OF SEQUENCES: 5256
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Human Genome Sciences, Inc.
;            STREET: 9410 Key West Avenue
;            CITY: Rockville
;            STATE: Maryland
;            COUNTRY: USA
;            ZIP: 20850
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;            COMPUTER: HP Vectra 486/33
;            OPERATING SYSTEM: MSDOS version 6.2
;            SOFTWARE: ASCII Text
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/956,171E
;            FILING DATE: 20-Oct-1997
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 60/009,861
;            FILING DATE: January 5, 1996
;            APPLICATION NUMBER: 08/781,986
;            FILING DATE: January 3, 1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Mark J. Hyman
;            REGISTRATION NUMBER: 46,789
;            REFERENCE/DOCKET NUMBER: PB248P1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (240) 314-1224
;            TELEFAX: (301) 309-8439
;   INFORMATION FOR SEQ ID NO: 5031:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 54 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 5031:
US-08-956-171E-5031

  Query Match             59.0%;  Score 11.8;  DB 4;  Length 54;
  Best Local Similarity   86.7%;  Pred. No. 5.2e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTT 16
              |||| |||||| |||
Db          7 GTAAAATCTATATTT 21


RESULT 39
US-08-956-171E-4977/c
; Sequence 4977, Application US/08956171E
; Patent No. 6593114
;   GENERAL INFORMATION:
;       APPLICANT: Charles Kunsch
;                  Gil H. Choi
;                  Patrick S. Dillon
;                  Craig A. Rosen
;                  Steven C. Barash
;                  Michael R. Fannon
;       TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;       NUMBER OF SEQUENCES: 5256
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Human Genome Sciences, Inc.
;            STREET: 9410 Key West Avenue
;            CITY: Rockville
;            STATE: Maryland
;            COUNTRY: USA
;            ZIP: 20850
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;            COMPUTER: HP Vectra 486/33
;            OPERATING SYSTEM: MSDOS version 6.2
;            SOFTWARE: ASCII Text
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/956,171E
;            FILING DATE: 20-Oct-1997
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 60/009,861
;            FILING DATE: January 5, 1996
;            APPLICATION NUMBER: 08/781,986
;            FILING DATE: January 3, 1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Mark J. Hyman
;            REGISTRATION NUMBER: 46,789
;            REFERENCE/DOCKET NUMBER: PB248P1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (240) 314-1224
;            TELEFAX: (301) 309-8439
;   INFORMATION FOR SEQ ID NO: 4977:
;       SEQUENCE CHARACTERISTICS:
```

```
;             LENGTH: 58 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 4977:
US-08-956-171E-4977

  Query Match             59.0%;  Score 11.8;  DB 4;  Length 58;
  Best Local Similarity   86.7%;  Pred. No. 5.2e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTT 16
              |||| |||||| |||
Db         25 GTAAAATCTATATTT 11


RESULT 40
US-09-462-941-30
; Sequence 30, Application US/09462941
; Patent No. 6608183
; GENERAL INFORMATION:
;  APPLICANT: Cox III, George N
;  APPLICANT:  Bolder Biotechnology, Inc.
;  TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins
;  FILE REFERENCE: 4152-1-PUS
;  CURRENT APPLICATION NUMBER: US/09/462,941
;  CURRENT FILING DATE:  2000-01-14
;  PRIOR APPLICATION NUMBER: 60/052,516
;  PRIOR FILING DATE: 1997-07-14
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
;   LENGTH: 65
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:PCR Primer
US-09-462-941-30

  Query Match             59.0%;  Score 11.8;  DB 4;  Length 65;
  Best Local Similarity   86.7%;  Pred. No. 5.3e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          6 CATCTATGTTTGGTT 20
              |||||||||| | ||
Db          2 CATCTATGTTCGTTT 16


RESULT 41
US-09-462-941-29
; Sequence 29, Application US/09462941
; Patent No. 6608183
; GENERAL INFORMATION:
;  APPLICANT: Cox III, George N
;  APPLICANT:  Bolder Biotechnology, Inc.
;  TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins
;  FILE REFERENCE: 4152-1-PUS
;  CURRENT APPLICATION NUMBER: US/09/462,941
;  CURRENT FILING DATE:  2000-01-14
;  PRIOR APPLICATION NUMBER: 60/052,516
;  PRIOR FILING DATE: 1997-07-14
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
;   LENGTH: 66
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:PCR Primer
US-09-462-941-29

  Query Match             59.0%;  Score 11.8;  DB 4;  Length 66;
  Best Local Similarity   86.7%;  Pred. No. 5.3e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          6 CATCTATGTTTGGTT 20
              |||||||||| | ||
Db         43 CATCTATGTTCGTTT 57


RESULT 42
US-08-398-617-5
; Sequence 5, Application US/08398617
; Patent No. 5747662
;  GENERAL INFORMATION:
;    APPLICANT:  Simmons, Laura C.
;    APPLICANT:  Yansura, Daniel G.
;    TITLE OF INVENTION:  Methods and Compositions for Secretion of Heterologous Proteir
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/398,617
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitts, Renee A.
;      REGISTRATION NUMBER:  35,136
;      REFERENCE/DOCKET NUMBER:  P889
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-1489
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  67 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-398-617-5

  Query Match             59.0%;  Score 11.8;  DB 1;  Length 67;
  Best Local Similarity   86.7%;  Pred. No. 5.3e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          6 CATCTATGTTTGGTT 20
              |||||||||| | ||
Db         37 CATCTATGTTCGTTT 51


RESULT 43
US-08-398-617-6/c
; Sequence 6, Application US/08398617
; Patent No. 5747662
;  GENERAL INFORMATION:
;    APPLICANT:  Simmons, Laura C.
;    APPLICANT:  Yansura, Daniel G.
;    TITLE OF INVENTION:  Methods and Compositions for Secretion of Heterologous Protein
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
```

```
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94080
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  patin (Genentech)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/398,617
;       FILING DATE:
;       CLASSIFICATION:  536
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Fitts, Renee A.
;       REGISTRATION NUMBER:  35,136
;       REFERENCE/DOCKET NUMBER:  P889
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415/225-1489
;       TELEFAX:  415/952-9881
;       TELEX:  910/371-7168
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  67 bases
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-398-617-6

  Query Match             59.0%;  Score 11.8;  DB 1;  Length 67;
  Best Local Similarity   86.7%;  Pred. No. 5.3e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          6 CATCTATGTTTGGTT 20
              ||||||||||  | ||
Db         35 CATCTATGTTCGTTT 21


RESULT 44
US-08-398-615-5
; Sequence 5, Application US/08398615
; Patent No. 5840523
;  GENERAL INFORMATION:
;    APPLICANT:  Simmons, Laura C.
;    APPLICANT:  Yansura, Daniel G.
;    TITLE OF INVENTION:  Methods and Compositions for Secretion of Heterologous Protein
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/398,615
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitts, Renee A.
;      REGISTRATION NUMBER:  35,136
;      REFERENCE/DOCKET NUMBER:  P889
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-1489
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  67 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-398-615-5

  Query Match             59.0%;  Score 11.8;  DB 2;  Length 67;
  Best Local Similarity   86.7%;  Pred. No. 5.3e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          6 CATCTATGTTTGGTT 20
              ||||||||||  | ||
Db         37 CATCTATGTTCGTTT 51


RESULT 45
US-08-398-615-6/c
; Sequence 6, Application US/08398615
; Patent No. 5840523
;  GENERAL INFORMATION:
;    APPLICANT:  Simmons, Laura C.
;    APPLICANT:  Yansura, Daniel G.
;    TITLE OF INVENTION:  Methods and Compositions for Secretion of Heterologous Protein
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/398,615
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitts, Renee A.
;      REGISTRATION NUMBER:  35,136
;      REFERENCE/DOCKET NUMBER:  P889
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-1489
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  67 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-398-615-6

  Query Match             59.0%;  Score 11.8;  DB 2;  Length 67;
  Best Local Similarity   86.7%;  Pred. No. 5.3e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          6 CATCTATGTTTGGTT 20
              ||||||||||  | ||
Db         35 CATCTATGTTCGTTT 21
```

Search completed: September 23, 2004, 16:44:22
Job time : 59 secs

This Page Blank (uspto)

OM nucleic - nucleic search, using sw model

Run on: September 23, 2004, 15:30:50 ; Search time 1335 Seconds
(without alignments)
447.373 Million cell updates/sec

Title: US-10-798-923A-36
Perfect score: 20
Sequence: 1 agtaacatctatgtttggtt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 375216

Minimum DB seq length: 0
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	14.2	71.0	42	29	BX203696	BX203696 Danio rer
	2	13.6	68.0	51	29	BX547845	BX547845 Arabidops
c	3	13.6	68.0	58	28	BZ287564	BZ287564 SALK_0209
c	4	13.6	68.0	73	28	BH811903	BH811903 SALK_0605
c	5	13.6	68.0	80	28	BH902197	BH902197 SALK_0914
c	6	13.4	67.0	46	28	AZ581263	AZ581263 1M0369N16
c	7	13.4	67.0	70	14	CD946310	CD946310 REL_89 Ge
c	8	13.4	67.0	75	9	AV836725	AV836725 AV836725
	9	13.2	66.0	31	28	AZ995788	AZ995788 2M0281D15
c	10	13.2	66.0	39	28	BH902806	BH902806 SALK_1009
c	11	13.2	66.0	46	28	AZ646477	AZ646477 1M0512G02
c	12	13.2	66.0	50	29	AL937476	AL937476 Arabidops
c	13	13.2	66.0	52	14	CB275049	CB275049 ku62b01.y
c	14	13.2	66.0	59	28	BZ384046	BZ384046 SALK_1349
c	15	13.2	66.0	79	28	BH613705	BH613705 SALK_0348
	16	13	65.0	76	9	AL682075	AL682075 AL682075
	17	12.8	64.0	44	28	BH907881	BH907881 SALK_0446
c	18	12.8	64.0	52	28	BH811539	BH811539 SALK_0590
c	19	12.8	64.0	64	9	AI930700	AI930700 sb38e11.y
	20	12.8	64.0	67	28	AZ783086	AZ783086 2M0024C04
c	21	12.8	64.0	70	29	TA26H12P	AL453461 T. brucei
c	22	12.8	64.0	73	29	BX286255	BX286255 Arabidops
	23	12.8	64.0	76	28	BZ592569	BZ592569 l(2)SH123
	24	12.8	64.0	77	28	AZ437102	AZ437102 1M0225K08
c	25	12.6	63.0	33	28	BH792403	BH792403 SALK_0641
	26	12.6	63.0	51	14	CA935569	CA935569 sau56b06.
c	27	12.6	63.0	52	9	AA065015	AA065015 zm12c05.r
c	28	12.6	63.0	58	9	AV518688	AV518688 AV518688
c	29	12.6	63.0	64	12	BI097343	BI097343 SWOv3MCAM
	30	12.6	63.0	66	28	AZ630914	AZ630914 1M0485A05
c	31	12.6	63.0	73	14	CD826259	CD826259 BN25.063E
	32	12.6	63.0	73	28	CC020433	CC020433 3591_1_19
	33	12.6	63.0	74	9	AI904585	AI904585 IL-BT062-
	34	12.6	63.0	74	9	AI904593	AI904593 IL-BT062-
c	35	12.6	63.0	74	29	DR17E12S	AL734254 Danio rer
	36	12.6	63.0	75	29	BX896687	BX896687 Arabidops
c	37	12.6	63.0	76	9	AI648529	AI648529 tz55a02.x
	38	12.6	63.0	79	28	BZ358128	BZ358128 SALK_1319
	39	12.4	62.0	49	28	AZ665591	AZ665591 1M0547D09
c	40	12.4	62.0	60	28	BH864251	BH864251 SALK_0956
	41	12.4	62.0	69	14	CD903663	CD903663 G356.110P
	42	12.4	62.0	77	29	CG602258	CG602258 OST275491
	43	12.4	62.0	79	14	CB911682	CB911682 VVD134E08
c	44	12.2	61.0	24	29	TA6A05P	AL452385 T. brucei
c	45	12.2	61.0	40	28	AZ774210	AZ774210 2M0003A19
c	46	12.2	61.0	43	28	BH847398	BH847398 SALK_0531
c	47	12.2	61.0	46	28	BH847200	BH847200 SALK_0445
c	48	12.2	61.0	46	28	BZ383285	BZ383285 SALK_1323
c	49	12.2	61.0	47	9	AW059738	AW059738 LE2a03.yg
	50	12.2	61.0	50	9	AL802723	AL802723 AL802723
c	51	12.2	61.0	51	28	BH861337	BH861337 SALK_0680
c	52	12.2	61.0	51	28	BZ353734	BZ353734 SALK_1220
	53	12.2	61.0	53	13	C00958	C00958 HUMGS000331
	54	12.2	61.0	55	29	BX533710	BX533710 Arabidops
	55	12.2	61.0	59	9	AU255386	AU255386 AU255386
c	56	12.2	61.0	59	14	CF326957	CF326957 NACL--01-
	57	12.2	61.0	59	14	CF326958	CF326958 NACL--01-
	58	12.2	61.0	62	12	BG145850	BG145850 uu99d04.y
c	59	12.2	61.0	62	28	BZ660481	BZ660481 SALK_0239
c	60	12.2	61.0	63	9	AA671132	AA671132 vn88a02.r
c	61	12.2	61.0	64	9	AI247077	AI247077 qw94b04.x
c	62	12.2	61.0	64	28	BH855346	BH855346 SALK_0859
c	63	12.2	61.0	65	29	AL938571	AL938571 Arabidops
c	64	12.2	61.0	66	28	AZ826358	AZ826358 2M0102G05
c	65	12.2	61.0	66	29	CC588867	CC588867 CH240_387
c	66	12.2	61.0	66	29	CG524790	CG524790 OST99133
	67	12.2	61.0	67	14	CD930683	CD930683 GR45.112A
c	68	12.2	61.0	69	28	BH811730	BH811730 SALK_0596
c	69	12.2	61.0	69	29	AL755239	AL755239 Arabidops
c	70	12.2	61.0	70	12	BG229757	BG229757 uu99d04.x
	71	12.2	61.0	70	29	AL945343	AL945343 Arabidops
c	72	12.2	61.0	75	29	CG649193	CG649193 OST403958
c	73	12.2	61.0	76	10	BE539012	BE539012 601059730
c	74	12.2	61.0	77	28	AZ454249	AZ454249 1M0256E07
c	75	12.2	61.0	77	29	CG647636	CG647636 OST397642
	76	12.2	61.0	79	10	BF016395	BF016395 uy40f01.y
	77	12	60.0	36	29	BX662610	BX662610 Arabidops

	78	12	60.0	37	28	AZ663202 AZ663202 1M0542P08
	79	12	60.0	40	9	AA907731 AA907731 oj93g12.s
c	80	12	60.0	43	28	AZ427738 AZ427738 1M0209I23
	81	12	60.0	45	28	CC325169 CC325169 XL921 Bay
	82	12	60.0	47	9	AJ235769 AJ235769 AJ235769
c	83	12	60.0	49	12	BM342828 BM342828 fw48e01.y
c	84	12	60.0	49	12	BM574161 BM574161 fx59h05.y
c	85	12	60.0	49	13	BQ077144 BQ077144 fz13b11.y
c	86	12	60.0	49	13	BQ615017 BQ615017 fab21c09.
	87	12	60.0	50	9	AU102956 AU102956 AU102956
c	88	12	60.0	50	12	BI705299 BI705299 fr56f02.y
c	89	12	60.0	50	12	BI840243 BI840243 fs71d07.y
c	90	12	60.0	50	13	BQ615523 BQ615523 fab27d09.
	91	12	60.0	50	28	BH866324 BH866324 SALK_1011
c	92	12	60.0	51	12	BM025452 BM025452 fs78g02.y
c	93	12	60.0	51	12	BM025460 BM025460 fs78h02.y
c	94	12	60.0	51	12	BM025463 BM025463 fs78h05.y
c	95	12	60.0	51	12	BM025535 BM025535 fs79h02.y
c	96	12	60.0	51	12	BM342608 BM342608 fw45h08.y
c	97	12	60.0	52	12	BI705096 BI705096 fr63e06.y
c	98	12	60.0	52	12	BM025311 BM025311 fs76g11.y
c	99	12	60.0	52	12	BM342786 BM342786 fw48a02.y
c	100	12	60.0	52	12	BM529195 BM529195 fy15b04.y
c	101	12	60.0	52	12	BM530824 BM530824 fy17c08.y
c	102	12	60.0	52	12	BM778857 BM778857 fx23e05.y
c	103	12	60.0	52	12	BM858211 BM858211 fx24e05.y
	104	12	60.0	53	9	AA721179 AA721179 nz71g07.s
c	105	12	60.0	53	12	BI702768 BI702768 fr60c11.y
c	106	12	60.0	53	12	BI702934 BI702934 fr66b03.y
c	107	12	60.0	53	12	BI702944 BI702944 fr66d08.y
c	108	12	60.0	53	12	BI702951 BI702951 fr66f05.y
c	109	12	60.0	53	12	BI705112 BI705112 fr63h12.y
c	110	12	60.0	53	12	BI705263 BI705263 fr55d12.y
c	111	12	60.0	53	12	BI705319 BI705319 fr57a05.y
c	112	12	60.0	53	12	BI708682 BI708682 fs57h11.y
c	113	12	60.0	53	12	BI709372 BI709372 fs63g06.y
c	114	12	60.0	53	12	BI839414 BI839414 fs67g10.y
c	115	12	60.0	53	12	BI840013 BI840013 fs68c01.y
c	116	12	60.0	53	12	BI840171 BI840171 fs70d04.y
c	117	12	60.0	53	12	BI840206 BI840206 fs70h01.y
c	118	12	60.0	53	12	BM025294 BM025294 fs76f02.y
c	119	12	60.0	53	12	BM025435 BM025435 fs78e05.y
c	120	12	60.0	53	12	BM025520 BM025520 fs79f07.y
c	121	12	60.0	53	12	BM025703 BM025703 fs82b03.y
c	122	12	60.0	53	12	BM181517 BM181517 fw43a11.y
c	123	12	60.0	53	12	BM183197 BM183197 fw31f01.y
c	124	12	60.0	53	12	BM186911 BM186911 fv79e09.y
c	125	12	60.0	53	12	BM187146 BM187146 fv82f08.y
c	126	12	60.0	53	12	BM341956 BM341956 fw54g11.y
c	127	12	60.0	53	12	BM341993 BM341993 fw55c05.y
c	128	12	60.0	53	12	BM342683 BM342683 fw46g11.y
c	129	12	60.0	53	12	BM343020 BM343020 fw50f04.y
c	130	12	60.0	53	12	BM778701 BM778701 fx21d12.y
c	131	12	60.0	54	12	BI702921 BI702921 fr65e05.y
c	132	12	60.0	54	12	BI703270 BI703270 fs47c09.y
c	133	12	60.0	54	12	BI703273 BI703273 fs47d06.y
c	134	12	60.0	54	12	BI703289 BI703289 fs47g08.y
c	135	12	60.0	54	12	BI705057 BI705057 fr62e11.y
c	136	12	60.0	54	12	BI708690 BI708690 fs58a10.y
c	137	12	60.0	54	12	BI708802 BI708802 fs59e11.y
c	138	12	60.0	54	12	BI709177 BI709177 fs60h03.y
c	139	12	60.0	54	12	BI709306 BI709306 fs62h07.y
c	140	12	60.0	54	12	BI839389 BI839389 fs67c12.y
c	141	12	60.0	54	12	BI839405 BI839405 fs67f05.y
c	142	12	60.0	54	12	BI840049 BI840049 fs68g01.y
c	143	12	60.0	54	12	BI840146 BI840146 fs70a07.y
c	144	12	60.0	54	12	BI840318 BI840318 fs72d09.y
c	145	12	60.0	54	12	BI840399 BI840399 fs73d11.y
c	146	12	60.0	54	12	BM025305 BM025305 fs76g04.y
c	147	12	60.0	54	12	BM025335 BM025335 fs77b09.y
c	148	12	60.0	54	12	BM025349 BM025349 fs77d04.y
c	149	12	60.0	54	12	BM025351 BM025351 fs77d06.y
c	150	12	60.0	54	12	BM025401 BM025401 fs78a08.y

ALIGNMENTS

```
RESULT 1
BX203696/c
LOCUS       BX203696                 42 bp    DNA     linear   GSS 29-JAN-2003
DEFINITION  Danio rerio genomic clone DKEY-223I20, genomic survey sequence.
ACCESSION   BX203696
VERSION     BX203696.1  GI:28035582
KEYWORDS    GSS.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Humphray,S.J., Huckle,E. and Durham,J.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Unpublished
COMMENT     This sequence was generated from the T7 end of BAC 223I20. 223I20
            is part of the Daniokey BAC Library created by R. Plasterk and N.V.
            Keygene. Further details:
            http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-223I20"
                     /tissue_type="Testis"
                     /note="vector pIndigoBAC-536"
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 29;  Length 42;
  Best Local Similarity   84.2%;  Pred. No. 2.3e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGTT 20
              ||||||  ||||||| |||
Db         29 GTAACAAGTATGTTTTGTT 11


RESULT 2
BX547845
LOCUS       BX547845                 51 bp    DNA     linear   GSS 02-JUL-2003
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-548H03-020587,
            genomic survey sequence.
ACCESSION   BX547845
VERSION     BX547845.1  GI:32440665
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H.
            and Weisshaar,B.
  TITLE     A pipeline for automated high-throughput generation of FSTs
            (flanking sequence tags) from Arabidopsis thaliana T-DNA
            transformed lines
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
  TITLE     A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
            for flanking sequence tag based reverse genetics
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 51)
  AUTHORS   Li,Y., Strizhov,N., Rosso,M. and Weisshaar,B.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUL-2003) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence is recovered from the left border of the T-DNA. It
            indicates an insertion close to or within gene At5g09430. The
            sequences are generated at the MPI for Plant Breeding Research in
            the context of the GABI-Kat project. GABI-Kat is part of the German
            Plant Genomics program designated 'GABI'. Information on line
            availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-548H03-020587"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161. The lines contain one or more T-DNA
                     insertions. The DNA fragment(s) resulting from the PCR
                     were directly sequenced to determine the genomic sequence
                     flanking the insertion. Sequences displaying significant
                     similarity to the A. thalaina nuclear genome sequence were
                     processed for submission. T-DNA derived sequences were
                     removed"
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 29;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 4.5e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGTT 20
              || |||| || || ||||||
Db         21 AGGAACACCTTTGATTGGTT 40


RESULT 3
BZ287564/c
LOCUS       BZ287564                 58 bp    DNA     linear   GSS 24-OCT-2002
DEFINITION  SALK_020942.48.40.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_020942.48.40.x, genomic
            survey sequence.
ACCESSION   BZ287564
VERSION     BZ287564.1  GI:24325637
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
            Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
            Shinn,P., Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within 300 bases of the 5' end of
            At1g10380.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_020942.48.40.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 28;  Length 58;
  Best Local Similarity   80.0%;  Pred. No. 4.6e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGTT 20
              | ||   ||||||||||||||
Db         35 ATTATTTTCTATGTTTGGTT 16


RESULT 4
BH811903/c
LOCUS       BH811903                 73 bp    DNA     linear   GSS 02-MAY-2002
DEFINITION  SALK_060502 Arabidopsis thaliana TDNA insertion lines Arabidopsis
            thaliana genomic clone SALK_060502, genomic survey sequence.
ACCESSION   BH811903
VERSION     BH811903.1  GI:20390358
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
            Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
            Shinn,P., Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within an annotated intron of At1g03687.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .73
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_060502"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 28;  Length 73;
  Best Local Similarity   80.0%;  Pred. No. 4.7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGTT 20
```

```
            ||| ||||||||||  | ||
Db         50 AGTGACATCTATGTAAGATT 31


RESULT 5
BH902197/c
LOCUS       BH902197                 80 bp    DNA     linear   GSS 04-SEP-2002
DEFINITION  SALK_091447.41.05.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_091447.41.05.x, genomic
            survey sequence.
ACCESSION   BH902197
VERSION     BH902197.1  GI:22713078
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 80)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
            Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
            Shinn,P., Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .80
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_091447.41.05.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 28;  Length 80;
  Best Local Similarity   80.0%;  Pred. No. 4.8e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGTT 20
              | ||||| |||| ||||| |
Db         67 AATAACAACTATATTTGGCT 48


RESULT 6
AZ581263/c
LOCUS       AZ581263                 46 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  1M0369N16R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0369N16 R, genomic survey sequence.
ACCESSION   AZ581263
VERSION     AZ581263.1  GI:11696100
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0369  row: N  column: 16
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 46.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0369N16"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 28;  Length 46;
  Best Local Similarity   93.3%;  Pred. No. 5.5e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTT 16
              || |||||||||||||
Db         16 GTTACATCTATGTTT 2


RESULT 7
CD946310/c
LOCUS       CD946310                 70 bp    mRNA    linear   EST 15-JUL-2003
DEFINITION  REL_89 GeneTag1 Zea mays cDNA, mRNA sequence.
ACCESSION   CD946310
VERSION     CD946310.1  GI:32794074
KEYWORDS    EST.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 70)
```

```
  AUTHORS   Genoplante.
  TITLE     Genoplante, a major partnership french program in plant genomics
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Genoplante
            Genoplante
            93, rue Henri Rochefort 91025 EVRY CEDEX France
            Tel: 33 1 69 47 54 00
            Fax: 33 1 69 47 54 10
            This sequence has been generated in the framework of the french
            plant genomics programme 'Genoplante' (http://www.genoplante.com
            and http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /cultivar="mixture"
                     /db_xref="taxon:4577"
                     /clone_lib="GeneTag1"
ORIGIN

  Query Match           67.0%;  Score 13.4;  DB 14;  Length 70;
  Best Local Similarity   93.3%;  Pred. No. 5.8e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 AACATCTATGTTTGG 18
              |||||| ||||||||
Db         60 AACATCCATGTTTGG 46


RESULT 8
AV836725/c
LOCUS       AV836725                 75 bp    mRNA    linear   EST 09-MAY-2002
DEFINITION  AV836725 K. Sato unpublished cDNA library: Hordeum vulgare subsp.
            vulgare seedling leaves second leaf stage Hordeum vulgare subsp.
            vulgare cDNA clone basd1a13, mRNA sequence.
ACCESSION   AV836725
VERSION     AV836725.1  GI:14528814
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare
  ORGANISM  Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 75)
  AUTHORS   Sato,K.
  TITLE     Barley EST sequencing project in NIG and Okayama Univ
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Kazuhiro Sato
            Research Institute for Bioresources
            Okayama University, Barley Germplasm Center
            Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
            Email: kazsato@rib.okayama-u.ac.jp,
            URL:http://www.rib.okayama-u.ac.jp/barley/
            database:http://www.shigen.nig.ac.jp/barley/Barley.html.
FEATURES             Location/Qualifiers
     source          1. .75
                     /organism="Hordeum vulgare subsp. vulgare"
                     /mol_type="mRNA"
                     /cultivar="Haruna Nijo"
                     /sub_species="vulgare"
                     /db_xref="taxon:112509"
                     /clone="basd1a13"
                     /tissue_type="seedling leaves"
                     /dev_stage="second leaf stage"
                     /clone_lib="K. Sato unpublished cDNA library: Hordeum
                     vulgare subsp. vulgare seedling leaves second leaf stage"
ORIGIN

  Query Match           67.0%;  Score 13.4;  DB 9;  Length 75;
  Best Local Similarity   87.5%;  Pred. No. 5.8e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGGTT 20
              |||| | |||||||||
Db         63 ACATGTNTGTTTGGTT 48


RESULT 9
AZ995788
LOCUS       AZ995788                 31 bp    DNA     linear   GSS 27-APR-2001
DEFINITION  2M0281D15R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
            clone UUGC2M0281D15 R, genomic survey sequence.
ACCESSION   AZ995788
VERSION     AZ995788.1  GI:13867015
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0281  row: D  column: 15
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 31.
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0281D15"
                     /sex="Female"
                     /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC2M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (female) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match           66.0%;  Score 13.2;  DB 28;  Length 31;
  Best Local Similarity   83.3%;  Pred. No. 6.4e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGG 18
```

```
              ||| |||| | |||||||
Db          1 AGTGACATGTGTGTTTGG 18


RESULT 10
BH902806/c
LOCUS       BH902806                39 bp    DNA     linear   GSS 04-SEP-2002
DEFINITION  SALK_100993.19.05.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_100993.19.05.x, genomic
            survey sequence.
ACCESSION   BH902806
VERSION     BH902806.1  GI:22713922
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 39)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
            Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
            Shinn,P., Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within an annotated intron of At5g12870.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .39
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_100993.19.05.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
ORIGIN

  Query Match           66.0%;  Score 13.2;  DB 28;  Length 39;
  Best Local Similarity  83.3%;  Pred. No. 6.6e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 TAACATCTATGTTTGGTT 20
              ||| |||||||| || ||
Db         27 TAATATCTATGTGTGTTT 10


RESULT 11
AZ646477/c
LOCUS       AZ646477                46 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  1M0512G02R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0512G02 R, genomic survey sequence.
ACCESSION   AZ646477
VERSION     AZ646477.1  GI:11776982
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0512  row: G  column: 02
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 46.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0512G02"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match           66.0%;  Score 13.2;  DB 28;  Length 46;
  Best Local Similarity  83.3%;  Pred. No. 6.7e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 TAACATCTATGTTTGGTT 20
              | |||||| ||| |||||
Db         39 TCACATCTCTGTGTGGTT 22


RESULT 12
AL937476/c
LOCUS       AL937476                50 bp    DNA     linear   GSS 23-OCT-2002
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-081A08-016179,
            genomic survey sequence.
ACCESSION   AL937476
VERSION     AL937476.1  GI:24369101
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
```

```
REFERENCE   1
  AUTHORS   Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H.
            and Weisshaar,B.
  TITLE     A pipeline for automated high-throughput generation of FSTs
            (flanking sequence tags) from Arabidopsis thaliana T-DNA
            transformed lines
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
  TITLE     A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
            for flanking sequence tag based reverse genetics
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 50)
  AUTHORS   Strizhov,N., Rosso,M., Li,Y. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-2002) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence is recovered from the left border of the T-DNA. It
            indicates an insertion within the locus defined by clone F14O13.
            The sequences are generated at the MPI for Plant Breeding Research
            in the context of the GABI-Kat project. GABI-Kat is part of the
            German Plant Genomics program designated 'GABI'. Information on
            line availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-081A08-016179"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161. The lines contain one or more T-DNA
                     insertions. The DNA fragment(s) resulting from the PCR
                     were directly sequenced to determine the genomic sequence
                     flanking the insertion. Sequences displaying significant
                     similarity to the A. thalaina nuclear genome sequence were
                     processed for submission. T-DNA derived sequences were
                     removed"
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 29;  Length 50;
  Best Local Similarity   83.3%;  Pred. No. 6.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 TAACATCTATGTTTGGTT 20
              ||| || || ||||||||
Db         34 TAAGATGTAGGTTTGGTT 17


RESULT 13
CB275049/c
LOCUS       CB275049                 52 bp    mRNA    linear   EST 24-FEB-2003
DEFINITION  ku62b01.y1 Strongyloides ratti PA female naive SL1 TOPO v2
            Strongyloides ratti cDNA 5', mRNA sequence.
ACCESSION   CB275049
VERSION     CB275049.1  GI:28504433
KEYWORDS    EST.
SOURCE      Strongyloides ratti
  ORGANISM  Strongyloides ratti
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Panagrolaimoidea; Strongyloididae; Strongyloides.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
            Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
            Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
            Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
            Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
            Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Claire Murphy and Dr. James McCarter
            at Washington University, St. Louis. This is a Oligo(dT)-SL1 PCR
            based library. cDNA PCR products of size >400 nucleotides
            containing SL1 on the 5' end and oligo(dT) on the 3' end were
            non-directionally cloned into pCRII-TOPO(Invitrogen) following the
            Topo TA cloning protocol. Parasitic adult females were collected
            from immunologically naive animals and provided by Dr. Mark Viney
            (Mark.Viney@bristol.ac.uk) of University of Bristol, Bristol, UK.
            Putative full length read
            The vector to vector length is 53
            Seq primer: SL1 primer.
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Strongyloides ratti"
                     /mol_type="mRNA"
                     /db_xref="taxon:34506"
                     /sex="female"
                     /dev_stage="Parasitic adult"
                     /lab_host="DH10B"
                     /clone_lib="Strongyloides ratti PA female naive SL1 TOPO
                     v2"
                     /note="Vector: pCRII-TOPO (Invitrogen); Site_1: EcoRI;
                     Site_2: EcoRI; The library was constructed by Claire
                     Murphy and Dr. James McCarter at Washington University,
                     St. Louis. This is a Oligo(dT)-SL1 PCR based library. cDNA
                     PCR products of size >400 nucleotides containing SL1 on
                     the 5' end and oligo(dT) on the 3' end were
                     non-directionally cloned into pCRII-TOPO(Invitrogen)
                     following the Topo TA cloning protocol. Parasitic adult
                     females were collected from immunologically naive animals
                     and provided by Dr. Mark Viney (Mark.Viney@bristol.ac.uk)
                     of University of Bristol, Bristol, UK."
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 14;  Length 52;
  Best Local Similarity   83.3%;  Pred. No. 6.9e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGT 19
              |||||||| || ||| ||
Db         40 GTAACATCCATTTTTAGT 23


RESULT 14
BZ384046/c
LOCUS       BZ384046                 59 bp    DNA     linear   GSS 26-NOV-2002
DEFINITION  SALK_134985.22.65.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_134985.22.65.x, genomic
            survey sequence.
ACCESSION   BZ384046
VERSION     BZ384046.1  GI:25480946
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
            Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
            Shinn,P., Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
```

```
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES              Location/Qualifiers
     source           1. .59
                      /organism="Arabidopsis thaliana"
                      /mol_type="genomic DNA"
                      /strain="Columbia 0"
                      /db_xref="taxon:3702"
                      /clone="SALK_134985.22.65.x"
                      /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                      /note="PCR was performed on Arabidopsis thaliana lines
                      each of which contains  one or more TDNA insertion
                      elements. The resultant fragment for each  line was
                      directly sequenced to determine the genomic sequence at
                      the  site of insertion. Details of the protocols used can
                      be found at  http://signal.salk.edu/tdna_protocols.html"
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 28;  Length 59;
  Best Local Similarity   83.3%;  Pred. No. 7e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGT 19
              ||||||||| ||  ||||
Db         25 GTAACATCTCTGAATGGT 8


RESULT 15
BH613705/c
LOCUS       BH613705                79 bp    DNA     linear   GSS 04-JAN-2002
DEFINITION  SALK_034815 Arabidopsis thaliana TDNA insertion lines Arabidopsis
            thaliana genomic clone SALK_034815, genomic survey sequence.
ACCESSION   BH613705
VERSION     BH613705.1  GI:18063073
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
            Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
            Shinn,P., Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES              Location/Qualifiers
     source           1. .79
                      /organism="Arabidopsis thaliana"
                      /mol_type="genomic DNA"
                      /strain="Columbia 0"
                      /db_xref="taxon:3702"
                      /clone="SALK_034815"
                      /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                      /note="PCR was performed on Arabidopsis thaliana lines
                      each of which contains  one or more TDNA insertion
                      elements. The resultant fragment for each  line was
                      directly sequenced to determine the genomic sequence at
                      the  site of insertion. Details of the protocols used can
                      be found at  http://signal.salk.edu/tdna_protocols.html"
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 28;  Length 79;
  Best Local Similarity   83.3%;  Pred. No. 7.3e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 TAACATCTATGTTTGGTT 20
              |||| ||| ||||||  ||
Db         29 TAACCTCTTTGTTTGTTT 12


RESULT 16
AL682075
LOCUS       AL682075                76 bp    mRNA    linear   EST 10-NOV-2003
DEFINITION  AL682075 XGC-gastrula Silurana tropicalis cDNA clone TGas058j17 5',
            mRNA sequence.
ACCESSION   AL682075
VERSION     AL682075.2  GI:38253903
KEYWORDS    EST.
SOURCE      Silurana tropicalis (western clawed frog)
  ORGANISM  Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Silurana.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
  TITLE     Sanger Xenopus tropicalis EST project 2001 (11_2003)
  JOURNAL   Unpublished (2003)
COMMENT     On Mar 18, 2002 this sequence version replaced gi:19538449.
            Contact: Taylor R
            Sanger Institute
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Aaron M. Zorn.
            cDNA was oligo dT primed from 5ug of poly A+ RNA from stages 10-13
            gastrulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with
            EcoRI at the 5' end and NotI at the 3' end.
            Vector: pCS107; Site_1: EcoRI; Site_2: NotI
            Host: Escherichia coli XL1-blue
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS_SEQUENCE_ID: TGas058j17.p1cSP6
            Sequencing primer: SP6.
FEATURES              Location/Qualifiers
     source           1. .76
                      /organism="Silurana tropicalis"
                      /mol_type="mRNA"
                      /db_xref="taxon:8364"
                      /clone="TGas058j17"
                      /dev_stage="gastrula (stages 10.5-12 mixed)"
                      /lab_host="Escherichia coli XL1-blue"
                      /clone_lib="XGC-gastrula"
                      /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
                      was oligo dT primed from 5ug of poly A+ RNA from stages
                      10-13 gastrulae.  EcoRI-NotI cut cDNA was then ligated
                      into pCS107 with EcoRI at the 5' end and NotI at the 3'
                      end."
ORIGIN

  Query Match             65.0%;  Score 13;  DB 9;  Length 76;
  Best Local Similarity   100.0%;  Pred. No. 8.9e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TAACATCTATGTT 15
              |||||||||||||
Db         64 TAACATCTATGTT 76
```

```
RESULT 17
BH907881
LOCUS       BH907881                 44 bp    DNA     linear   GSS 04-SEP-2002
DEFINITION  SALK_044616.22.70.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_044616.22.70.x, genomic
            survey sequence.
ACCESSION   BH907881
VERSION     BH907881.1  GI:22720814
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
            Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
            Shinn,P., Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_044616.22.70.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 28;  Length 44;
  Best Local Similarity   87.5%;  Pred. No. 1e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGGTT 20
              | | ||||||||||||
Db         11 AAAACTATGTTTGGTT 26


RESULT 18
BH811539/c
LOCUS       BH811539                 52 bp    DNA     linear   GSS 02-MAY-2002
DEFINITION  SALK_059062 Arabidopsis thaliana TDNA insertion lines Arabidopsis
            thaliana genomic clone SALK_059062, genomic survey sequence.
ACCESSION   BH811539
VERSION     BH811539.1  GI:20389994
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
            Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
            Shinn,P., Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within an annotated intron of At5g54840.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_059062"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 28;  Length 52;
  Best Local Similarity   87.5%;  Pred. No. 1e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTT 16
              | ||| ||||||||||
Db         26 ATTAATATCTATGTTT 11


RESULT 19
AI930700/c
LOCUS       AI930700                 64 bp    mRNA    linear   EST 30-NOV-2001
DEFINITION  sb38e11.y1 Gm-c1013 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1013-357 5' similar to SW:G3PC_DIACA P34921 GLYCERALDEHYDE
            3-PHOSPHATE DEHYDROGENASE, CYTOSOLIC ;, mRNA sequence.
ACCESSION   AI930700
VERSION     AI930700.1  GI:5666664
KEYWORDS    EST.
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 64)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
            Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
            Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
            Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
            Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand This clone is
            available through: ResGen, Invitrogen Corp. 2130 South Memorial
            Parkway Huntsville, AL 35801 For further information call:
```

(800)-533-4363 or contact via email: ccu@resgen.com
Insert Length: 1417 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 1.

FEATURES Location/Qualifiers
source 1. .64
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1013-357"
/tissue_type="Whole seedlings, 2-3 week old seedlings, greenhouse grown"
/lab_host="XL10-Gold"
/clone_lib="Gm-c1013"
/note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2: XhoI; This cDNA library was constructed from mRNA isolated from whole seedlings of 2-3 week old greenhouse grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR cDNA library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelding."

ORIGIN

Query Match 64.0%; Score 12.8; DB 9; Length 64;
Best Local Similarity 87.5%; Pred. No. 1.1e+05;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          4 AACATCTATGTTTGGT 19
              ||||| |||||| |||
Db         27 AACATGTATGTTAGGT 12
```

RESULT 20
AZ783086
LOCUS AZ783086 67 bp DNA linear GSS 16-FEB-2001
DEFINITION 2M0024C04R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0024C04 R, genomic survey sequence.
ACCESSION AZ783086
VERSION AZ783086.1 GI:12917459
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 67)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0024 row: C column: 04
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 67.

FEATURES Location/Qualifiers
source 1. .67
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0024C04"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 64.0%; Score 12.8; DB 28; Length 67;
Best Local Similarity 87.5%; Pred. No. 1.1e+05;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 AGTAACATCTATGTTT 16
              ||| ||| ||||||||
Db         52 AGTTACACCTATGTTT 67
```

RESULT 21
TA26H12P/c
LOCUS TA26H12P 70 bp DNA linear GSS 13-DEC-2000
DEFINITION T. brucei sheared genomic DNA clone 26h12, forward sequence, genomic survey sequence.
ACCESSION AL453461
VERSION AL453461.1 GI:11850973
KEYWORDS GSS.
SOURCE Trypanosoma brucei
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
REFERENCE 1 (bases 1 to 70)
AUTHORS Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R., Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L., Melville,S.E., Rajandream,M.A. and Barrell,B.G.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nh1@sanger.ac.uk
COMMENT Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (
4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/T_brucei/.

FEATURES Location/Qualifiers

```
     source          1. .70
                     /organism="Trypanosoma brucei"
                     /mol_type="genomic DNA"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="26h12"
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 29;  Length 70;
  Best Local Similarity   87.5%;  Pred. No. 1.1e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 AACATCTATGTTTGGT 19
              ||||||||||| | ||
Db         44 AACATCTATGTGTTGT 29


RESULT 22
BX286255/c
LOCUS       BX286255                  73 bp    DNA     linear   GSS 07-MAR-2003
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-389F02-018268,
            genomic survey sequence.
ACCESSION   BX286255
VERSION     BX286255.1  GI:28885251
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H.
            and Weisshaar,B.
  TITLE     A pipeline for automated high-throughput generation of FSTs
            (flanking sequence tags) from Arabidopsis thaliana T-DNA
            transformed lines
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
  TITLE     A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
            for flanking sequence tag based reverse genetics
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 73)
  AUTHORS   Strizhov,N., Li,Y., Rosso,M. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-MAR-2003) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence is recovered from the left border of the T-DNA. It
            indicates an insertion within the locus defined by clone F5E19. The
            sequences are generated at the MPI for Plant Breeding Research in
            the context of the GABI-Kat project. GABI-Kat is part of the German
            Plant Genomics program designated 'GABI'. Information on line
            availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .73
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-389F02-018268"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161. The lines contain one or more T-DNA
                     insertions. The DNA fragment(s) resulting from the PCR
                     were directly sequenced to determine the genomic sequence
                     flanking the insertion. Sequences displaying significant
                     similarity to the A. thalaina nuclear genome sequence were
                     processed for submission. T-DNA derived sequences were
                     removed"
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 29;  Length 73;
  Best Local Similarity   87.5%;  Pred. No. 1.1e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGGTT 20
              |||||| | |||||||
Db         47 ACATCTGTATTTGGTT 32


RESULT 23
BZ592569
LOCUS       BZ592569                  76 bp    DNA     linear   GSS 24-FEB-2003
DEFINITION  l(2)SH1233 P-element-insertion screen Drosophila melanogaster
            genomic clone l(2)SH1233 5', genomic survey sequence.
ACCESSION   BZ592569
VERSION     BZ592569.1  GI:27414547
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Oh,S.W., Kingsley,T., Shin,H., Zheng,Z., Chen.H.W., Chen,X.,
            Wang,H., Ruan,P., Moody,M. and Hou,S.X.
  TITLE     A P-element-insertion screen identified mutations in 450 novel
            genes in Drosophila
  JOURNAL   Genetics 163 (1), 195-201 (2003)
  MEDLINE   22474228
   PUBMED   12586707
COMMENT     Contact: Suwan Oh
            Lab. of Immunobiology
            NCI-FCRDC
            1050 Boyles St., Frederick, MD 21702, USA
            Tel: 1-301-846-7314
            Fax: 1-301-846-6145
            Email: ohsuwan@mail.ncifcrf.gov
            Class: unknown
            High quality sequence stop: 76.
FEATURES             Location/Qualifiers
     source          1. .76
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="l(2)SH1233"
                     /clone_lib="P-element-insertion screen"
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 28;  Length 76;
  Best Local Similarity   87.5%;  Pred. No. 1.1e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTT 16
              || |||||| ||||||
Db         28 AGAAACATCGATGTTT 43


RESULT 24
AZ437102
LOCUS       AZ437102                  77 bp    DNA     linear   GSS 03-OCT-2000
DEFINITION  1M0225K08F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0225K08 F, genomic survey sequence.
ACCESSION   AZ437102
VERSION     AZ437102.1  GI:10561115
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 77)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
```

```
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0225  row: K  column: 08
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 77.
FEATURES             Location/Qualifiers
     source          1. .77
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0225K08"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 28;  Length 77;
  Best Local Similarity   87.5%;  Pred. No. 1.1e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 TAACATCTATGTTTGG 18
              | |||||||| |||||
Db         53 TGACATCTATTTTTGG 68


RESULT 25
BH792403/c
LOCUS       BH792403                 33 bp    DNA     linear   GSS 02-APR-2002
DEFINITION  SALK_064174.25.60.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_064174.25.60.x, genomic
            survey sequence.
ACCESSION   BH792403
VERSION     BH792403.1  GI:19889138
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 33)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
            Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
            Shinn,P., Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within 300 bases of the 3' end of
            At5g02890.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .33
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_064174.25.60.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 28;  Length 33;
  Best Local Similarity   78.9%;  Pred. No. 1.2e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              ||| ||||  ||| |||||
Db         25 AGTGACATGCATGGTTGGT 7


RESULT 26
CA935569
LOCUS       CA935569                 51 bp    mRNA    linear   EST 30-DEC-2002
DEFINITION  sau56b06.y1 Gm-c1071 Glycine max cDNA clone SOYBEAN CLONE ID:
            Gm-c1071-4692 5', mRNA sequence.
ACCESSION   CA935569
VERSION     CA935569.1  GI:27424049
KEYWORDS    EST.
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
            Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
            Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
            Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
            Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
```

call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com
Putative full length read
vector to vector length is 52
Seq primer: -40RP from Gibco.

FEATURES Location/Qualifiers
source 1. .51
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-c1071-4692"
/tissue_type="immature pods (~2cm long) of greenhouse grown plants"
/lab_host="DH10B"
/clone_lib="Gm-c1071"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; The cDNA library was constructed from mRNA isolated from immature pods (approximately 2cm long) of greenhouse grown plants. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly(dT) sequence with a NotI restrictions site. SalI linkers adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Anu Khanna at the University of Illinios at Urbana-Champaign. email: l-vodkin@uiuc.edu"

ORIGIN

Query Match 63.0%; Score 12.6; DB 14; Length 51;
Best Local Similarity 78.9%; Pred. No. 1.3e+05;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 GTAACATCTATGTTTGGTT 20
|||| || |||||| |||
Db 21 GTAAAATATATGTTGTGTT 39

RESULT 27
AA065015/c
LOCUS AA065015 52 bp mRNA linear EST 23-DEC-1997
DEFINITION zm12c05.r1 Stratagene pancreas (#937208) Homo sapiens cDNA clone IMAGE:525416 5' similar to TR:G545018 G545018 BRG1=BRAHMA HOMOLOG ;, mRNA sequence.
ACCESSION AA065015
VERSION AA065015.1 GI:1558631
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 52)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
PUBMED 8889549
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
WARNING: There is evidence that suggests that the 384-well parent plate of this clone contains both human and mouse derived clones. Thus, the origin of this clone is uncertain. This caution should be kept in mind should you use this clone.

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 1.

FEATURES Location/Qualifiers
source 1. .52
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3916845"
/db_xref="taxon:9606"
/clone="IMAGE:525416"
/lab_host="SOLR cells (kanamycin resistant)"
/clone_lib="Stratagene pancreas (#937208)"
/note="Organ: pancreas; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Pancreatic adenocarcinoma cell line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"

ORIGIN

Query Match 63.0%; Score 12.6; DB 9; Length 52;
Best Local Similarity 78.9%; Pred. No. 1.3e+05;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 GTAACATCTATGTTTGGTT 20
|||||||| | |||||
Db 44 GGAACATCTAGGGGTGGTT 26

RESULT 28
AV518688/c
LOCUS AV518688 58 bp mRNA linear EST 29-SEP-2000
DEFINITION AV518688 Arabidopsis thaliana aboveground organs two to six-week old Arabidopsis thaliana cDNA clone APD36a04F 3', mRNA sequence.
ACCESSION AV518688
VERSION AV518688.2 GI:10423370
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 58)
AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
JOURNAL DNA Res. 7 (3), 175-180 (2000)
MEDLINE 20363093
PUBMED 10907847
COMMENT On Jun 23, 2000 this sequence version replaced gi:8678215.
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

FEATURES Location/Qualifiers
source 1. .58
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="APD36a04F"
/tissue_type="aboveground organs"
/dev_stage="two to six-week old"
/clone_lib="Arabidopsis thaliana aboveground organs two to

six-week old"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"
ORIGIN

Query Match 63.0%; Score 12.6; DB 9; Length 58;
Best Local Similarity 78.9%; Pred. No. 1.3e+05;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 GTAACATCTATGTTTGGTT 20
||||| | |||| ||| ||
Db 40 GTAACTTGTATGATTGATT 22

RESULT 29
BI097343/c
LOCUS BI097343 64 bp mRNA linear EST 25-JUN-2001
DEFINITION SWOv3MCAM61F08SK Onchocerca volvulus molting L3 larva cDNA (SL96MLW-OvmL3) Onchocerca volvulus cDNA clone SWOv3MCAM61F08 5', mRNA sequence.
ACCESSION BI097343
VERSION BI097343.1 GI:14549000
KEYWORDS EST.
SOURCE Onchocerca volvulus
ORGANISM Onchocerca volvulus
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea; Onchocercidae; Onchocerca.
REFERENCE 1 (bases 1 to 64)
AUTHORS Williams,S.A., Lizotte-Waniewski,M., Laney,S. and Lustigman,S.
TITLE Genes expressed in molting L3 larvae of Onchocerca volvulus
JOURNAL Unpublished (1997)
COMMENT Contact: Steven A. Williams
Molecular Parasitology
Smith College Department of Biological Sciences
Department of Biological Sciences, Clark Science Center, Smith College, Northampton, MA, 01063, USA
Tel: 4135853826
Fax: 4135853786
Email: genome@smith.edu
Seq primer: pBluescript SK.
FEATURES Location/Qualifiers
source 1. .64
/organism="Onchocerca volvulus"
/mol_type="mRNA"
/strain="Kumba, Cameroons"
/db_xref="taxon:6282"
/clone="SWOv3MCAM61F08"
/dev_stage="molting L3"
/lab_host="XL1-Blue MRF'"
/clone_lib="Onchocerca volvulus molting L3 larva cDNA (SL96MLW-OvmL3)"
/note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2: Xho I; Filarial nematode parasite of humans. Third-stage larvae, L3, were isolated from infected black flies in Cameroon (forest strain). The L3 were cultured in 20% FCS in IMDM+ NCTC 135 and collected after day 1, 2, or 3 in culture. L3 of O. volvulus molt to fourth-stage larvae by day 5 in culture. mRNA was isolated from approximately 6000 molting larvae (mL3),2000 larvae from day 1, 2 or 3 in culture, and converted to double-stranded cDNA using reverse transcriptase and oligo(dT) followed by RNase H and DNA pol I. The library was constructed in the lambda Uni-Zap XR vector and has 1 x 10E6 independent recombinants and the average insert size is ~1200 bp. The library was constructed by Sara Lustigman and Michelle Lizotte-Waniewski in the Laboratory of Dr. S. A. Williams. The library is available from Dr. Sara Lustigman (email: slustigm@nybc.org)."
ORIGIN

Query Match 63.0%; Score 12.6; DB 12; Length 64;
Best Local Similarity 78.9%; Pred. No. 1.3e+05;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AGTAACATCTATGTTTGGT 19
|| | ||||| |||||| |
Db 48 AGAAGCATCTGTGTTTGAT 30

RESULT 30
AZ630914
LOCUS AZ630914 66 bp DNA linear GSS 13-DEC-2000
DEFINITION 1M0485A05F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0485A05 F, genomic survey sequence.
ACCESSION AZ630914
VERSION AZ630914.1 GI:11753200
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 66)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0485 row: A column: 05
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 66.
FEATURES Location/Qualifiers
source 1. .66
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0485A05"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
ORIGIN

Query Match 63.0%; Score 12.6; DB 28; Length 66;
Best Local Similarity 78.9%; Pred. No. 1.3e+05;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 AGTAACATCTATGTTTGGT 19
              |||| | |||  |||||||
Db         43 AGTAGCTTCTGGGTTTGGT 61


RESULT 31
CD826259/c
LOCUS       CD826259                 73 bp    mRNA    linear   EST 10-JUL-2003
DEFINITION  BN25.063E12F020118 BN25 Brassica napus cDNA clone BN25063E12, mRNA
            sequence.
ACCESSION   CD826259
VERSION     CD826259.1  GI:32508199
KEYWORDS    EST.
SOURCE      Brassica napus (rape)
  ORGANISM  Brassica napus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   Genoplante.
  TITLE     Genoplante, a major partnership french program in plant genomics
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Genoplante
            Genoplante
            93, rue Henri Rochefort 91025 EVRY CEDEX France
            Tel: 33 1 69 47 54 00
            Fax: 33 1 69 47 54 10
            This sequence has been generated in the framework of the french
            plant genomics programme 'Genoplante' (http://www.genoplante.com
            and http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .73
                     /organism="Brassica napus"
                     /mol_type="mRNA"
                     /cultivar="Jet neuf"
                     /db_xref="taxon:3708"
                     /clone="BN25063E12"
                     /tissue_type="seed"
                     /clone_lib="BN25"
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 14;  Length 73;
  Best Local Similarity   78.9%;  Pred. No. 1.4e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGTT 20
              | ||||| |||||||  ||
Db         66 GCAACATGTATGTTTTTTT 48


RESULT 32
CC020433
LOCUS       CC020433                 73 bp    DNA     linear   GSS 01-APR-2003
DEFINITION  3591_1_19_1_H05.2EL_y_1 3591 - RescueMu Grid P Zea mays genomic,
            genomic survey sequence.
ACCESSION   CC020433
VERSION     CC020433.1  GI:29434506
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Possible ligation site of ends cut by 2 different endonucleases.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 3591_1_19_1  row: 12
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .73
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="mixed background W23/A188/B73/K55"
                     /db_xref="taxon:4577"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="3591 - RescueMu Grid P"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.' Grid P was grown at Molokai in 2002. DNA was
                     extracted from leaf strips, double digested using BamHI
                     and BglII, and ligated to form circular plasmids. DH10B
                     cells were transformed and then screened on LB plates with
                     ampicillin."
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 28;  Length 73;
  Best Local Similarity   78.9%;  Pred. No. 1.4e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              ||||||   | ||||||||
Db         36 AGTAACTGTTCTGTTTGGT 54


RESULT 33
AI904585
LOCUS       AI904585                 74 bp    mRNA    linear   EST 30-MAR-2000
DEFINITION  IL-BT062-191298-010_1 BT062 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AI904585
VERSION     AI904585.1  GI:6494972
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 74)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/seq/gethtml.pl?t1=IL&t2=IL-BT062-010_1.ht
```

```
          ml&t3=191298&t4=1)
          Seq primer: puc 18 forward.
FEATURES             Location/Qualifiers
     source          1. .74
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /sex="female"
                     /dev_stage="Adult"
                     /clone_lib="BT062"
                     /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No.
                     196,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 9;  Length 74;
  Best Local Similarity   78.9%;  Pred. No. 1.4e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              || |||| || | ||||||
Db         12 AGAAACAGCTCTCTTTGGT 30


RESULT 34
AI904593
LOCUS       AI904593                 74 bp    mRNA    linear   EST 30-MAR-2000
DEFINITION  IL-BT062-311298-011 BT062 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AI904593
VERSION     AI904593.1  GI:6494980
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 74)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/seq/gethtml.pl?t1=IL&t2=IL-BT062-011.html
            &t3=311298&t4=1)
            Seq primer: puc 18 forward.
FEATURES             Location/Qualifiers
     source          1. .74
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /sex="female"
                     /dev_stage="Adult"
                     /clone_lib="BT062"
                     /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No.
                     196,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 9;  Length 74;
  Best Local Similarity   78.9%;  Pred. No. 1.4e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              || |||| || | ||||||
Db         12 AGAAACAGCTCTCTTTGGT 30


RESULT 35
DR17E12S/c
LOCUS       DR17E12S                 74 bp    DNA     linear   GSS 27-NOV-2002
DEFINITION  Danio rerio genomic clone DKEY-17E12, genomic survey sequence.
ACCESSION   AL734254
VERSION     AL734254.1  GI:21342336
KEYWORDS    GSS.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 74)
  AUTHORS   Humphray,S.J., Huckle,E. and Hunt,S.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUN-2002) The Sanger Institute, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
            humquery@sanger.ac.uk Unpublished
COMMENT     This sequence was generated from the SP6 end of BAC 17E12. 17E12 is
            part of the Daniokey BAC Library created by R. Plasterk and N.V.
            Keygene.
            Further details: http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES             Location/Qualifiers
     source          1. .74
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-17E12"
                     /tissue_type="Testis"
                     /note="vector pIndigoBAC-536"
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 29;  Length 74;
  Best Local Similarity   78.9%;  Pred. No. 1.4e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGTT 20
              | || |  ||||||||||
Db         59 GAAAAAAATATGTTTGGTT 41


RESULT 36
BX896687
LOCUS       BX896687                 75 bp    DNA     linear   GSS 15-DEC-2003
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-759A06-023802,
            genomic survey sequence.
ACCESSION   BX896687
VERSION     BX896687.1  GI:39929182
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H.
```

```
            and Weisshaar,B.
  TITLE     A pipeline for automated high-throughput generation of FSTs
            (flanking sequence tags) from Arabidopsis thaliana T-DNA
            transformed lines
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
  TITLE     A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
            for flanking sequence tag based reverse genetics
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 75)
  AUTHORS   Rosso,M., Li,Y., Strizhov,N. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-DEC-2003) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence is recovered from the left border of the T-DNA. It
            indicates an insertion within the locus defined by clone t3f17. The
            sequences are generated at the MPI for Plant Breeding Research in
            the context of the GABI-Kat project. GABI-Kat is part of the German
            Plant Genomics program designated 'GABI'. Information on line
            availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .75
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-759A06-023802"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161. The lines contain one or more T-DNA
                     insertions. The DNA fragment(s) resulting from the PCR
                     were directly sequenced to determine the genomic sequence
                     flanking the insertion. Sequences displaying significant
                     similarity to the A. thalaina nuclear genome sequence were
                     processed for submission. T-DNA derived sequences were
                     removed"
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 29;  Length 75;
  Best Local Similarity   78.9%;  Pred. No. 1.4e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGTT 20
              |||||| || | |||| ||
Db         21 GTAACAACTTTATTTGTTT 39


RESULT 37
AI648529/c
LOCUS       AI648529                 76 bp    mRNA    linear   EST 16-DEC-1999
DEFINITION  tz55a02.x1 NCI_CGAP_Ov35 Homo sapiens cDNA clone IMAGE:2292458 3'
            similar to gb:M60278 HEPARIN-BINDING EGF-LIKE GROWTH FACTOR
            PRECURSOR (HUMAN);contains MSR1.t3 TAR1 repetitive element ;, mRNA
            sequence.
ACCESSION   AI648529
VERSION     AI648529.1  GI:4729363
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 254   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .76
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2292458"
                     /tissue_type="tumor, 5 pooled (see description)"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Ov35"
                     /note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; This library represents the normalized
                     version of NCI_CGAP_Ov23. Cloned unidirectionally.
                     Primer: Oligo dT.  Average insert size 0.86 kb. Tumor
                     types include: mixed Mullerian tumor, papillary serous,
                     clear cell, spindle cell.  All are primary tumors,
                     metastasis positive. Constructed by Life Technologies."
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 9;  Length 76;
  Best Local Similarity   78.9%;  Pred. No. 1.4e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              | ||||| |||| |||| |
Db         64 AATAACAACTATTTTGTT 46


RESULT 38
BZ358128
LOCUS       BZ358128                 79 bp    DNA     linear   GSS 14-NOV-2002
DEFINITION  SALK_131965.28.05.n Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_131965.28.05.n, genomic
            survey sequence.
ACCESSION   BZ358128
VERSION     BZ358128.1  GI:24950291
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
            Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
            Shinn,P., Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .79
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
```

/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_131965.28.05.n"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN

Query Match 63.0%; Score 12.6; DB 28; Length 79;
Best Local Similarity 78.9%; Pred. No. 1.4e+05;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AGTAACATCTATGTTTGGT 19
||||| || ||||||||
Db 13 AGTAAGATAATTGTTTGGT 31

RESULT 39
AZ665591
LOCUS AZ665591 49 bp DNA linear GSS 14-DEC-2000
DEFINITION 1M0547D09F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0547D09 F, genomic survey sequence.
ACCESSION AZ665591
VERSION AZ665591.1 GI:11802737
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 49)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0547 row: D column: 09
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 49.
FEATURES Location/Qualifiers
source 1. .49
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0547D09"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 62.0%; Score 12.4; DB 28; Length 49;
Best Local Similarity 92.9%; Pred. No. 1.6e+05;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 AACATCTATGTTTG 17
|||||| |||||||
Db 11 AACATCCATGTTTG 24

RESULT 40
BH864251/c
LOCUS BH864251 60 bp DNA linear GSS 05-AUG-2002
DEFINITION SALK_095643 Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK_095643, genomic survey sequence.
ACCESSION BH864251
VERSION BH864251.1 GI:22100149
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 60)
AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R.
TITLE A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
JOURNAL Unpublished (2001)
COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of TDNA. This sequence lies within an annotated intron of At1g32240.
Class: TDNA tagged.
FEATURES Location/Qualifiers
source 1. .60
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_095643"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN

Query Match 62.0%; Score 12.4; DB 28; Length 60;
Best Local Similarity 92.9%; Pred. No. 1.6e+05;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 ATCTATGTTTGGTT 20
|||||||||||| |

```
Db         37 ATCTATGTTTGGCT 24


RESULT 41
CD903663
LOCUS       CD903663                 69 bp    mRNA    linear   EST 14-JUL-2003
DEFINITION  G356.110P16F010919 G356 Triticum aestivum cDNA clone G356110P16,
            mRNA sequence.
ACCESSION   CD903663
VERSION     CD903663.1  GI:32677991
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 69)
  AUTHORS   Genoplante.
  TITLE     Genoplante, a major partnership french program in plant genomics
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Genoplante
            Genoplante
            93, rue Henri Rochefort 91025 EVRY CEDEX France
            Tel: 33 1 69 47 54 00
            Fax: 33 1 69 47 54 10
            This sequence has been generated in the framework of the french
            plant genomics programme 'Genoplante' (http://www.genoplante.com
            and http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .69
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="recital"
                     /db_xref="taxon:4565"
                     /clone="G356110P16"
                     /tissue_type="grain (356 degrees per day after
                     pollination)"
                     /clone_lib="G356"
ORIGIN

  Query Match             62.0%;  Score 12.4;  DB 14;  Length 69;
  Best Local Similarity   92.9%;  Pred. No. 1.7e+05;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGG 18
              |||||||||||| |
Db         30 ACATCTATGTTTTG 43


RESULT 42
CG602258
LOCUS       CG602258                 77 bp    DNA     linear   GSS 02-OCT-2003
DEFINITION  OST275491 Mus musculus 129Sv/Ev Mus musculus genomic clone
            OST275491, genomic survey sequence.
ACCESSION   CG602258
VERSION     CG602258.1  GI:37421753
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 77)
  AUTHORS   Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
            Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
            Friddle,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
            Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D.,
            Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
            Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
            Zhu,Q., Person,C. and Sands,A.T.
  TITLE     Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
            screen to identify potential targets for therapeutic intervention
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT     Contact: Zambrowicz BP
            OmniBank
            Lexicon Genetics Incorporated
            4000 Research Forest Drive, The Woodlands, TX 77381, USA
            Email: materials@lexgen.com
            Gene trap sequence tag generated by 3' RACE from mouse ES cells as
            described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
            Class: Gene Trap.
FEATURES             Location/Qualifiers
     source          1. .77
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129SV/EV"
                     /db_xref="taxon:10090"
                     /clone="OST275491"
                     /cell_type="embryonic stem cell"
                     /clone_lib="Mus musculus 129Sv/Ev"
ORIGIN

  Query Match             62.0%;  Score 12.4;  DB 29;  Length 77;
  Best Local Similarity   86.7%;  Pred. No. 1.7e+05;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 TAACATCTATGTTTG 17
              ||| |||||| ||||
Db         40 TAANATCTATTTTTG 54


RESULT 43
CB911682
LOCUS       CB911682                 79 bp    mRNA    linear   EST 25-APR-2003
DEFINITION  VVD134E08_373255 An expressed sequence tag database for abiotic
            stressed berries of Vitis vinifera var. Chardonnay Vitis vinifera
            cDNA clone VVD134E08 5, mRNA sequence.
ACCESSION   CB911682
VERSION     CB911682.1  GI:30126343
KEYWORDS    EST.
SOURCE      Vitis vinifera
  ORGANISM  Vitis vinifera
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; Vitaceae; Vitis.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   Cushman,J.C.
  TITLE     An expressed sequence tag database for abiotic stressed berries of
            Vitis vinifera var. Chardonnay
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Cushman JC
            Department of Biochemistry
            University of Nevada
            MS200, Reno, NV 89557-0014, USA
            Tel: 775-784-1918
            Fax: 775-784-1650
            Email: jcushman@unr.edu
            PCR PRimers
            FORWARD: T3 20mer
            BACKWARD: T7 21mer (backward)
            Plate: 134  row: E  column: 08
            Seq primer: T3 20mer
            High quality sequence stop: 79.
FEATURES             Location/Qualifiers
     source          1. .79
                     /organism="Vitis vinifera"
                     /mol_type="mRNA"
                     /db_xref="taxon:29760"
                     /clone="VVD134E08"
                     /tissue_type="berries"
                     /dev_stage="mixed; 8, 9, 11, 13, 15, 16 weeks daf"
                     /clone_lib="An expressed sequence tag database for abiotic
                     stressed berries of Vitis vinifera var. Chardonnay"
                     /note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI"
ORIGIN
```

```
  Query Match             62.0%;  Score 12.4;  DB 14;  Length 79;
  Best Local Similarity   92.9%;  Pred. No. 1.7e+05;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 ATCTATGTTTGGTT 20
              |||| |||||||||
Db         12 ATCTTTGTTTGGTT 25


RESULT 44
TA6A05P/c
LOCUS       TA6A05P                   24 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 6a05, forward sequence, genomic
            survey sequence.
ACCESSION   AL452385
VERSION     AL452385.1  GI:11857848
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="Trypanosoma brucei"
                     /mol_type="genomic DNA"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="6a05"
ORIGIN

  Query Match             61.0%;  Score 12.2;  DB 29;  Length 24;
  Best Local Similarity   82.4%;  Pred. No. 1.8e+05;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTG 17
              ||||||||   ||||||
Db         24 AGTAACATAACTGTTTG 8


RESULT 45
AZ774210/c
LOCUS       AZ774210                  40 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  2M0003A19F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0003A19 F, genomic survey sequence.
ACCESSION   AZ774210
VERSION     AZ774210.1  GI:12899399
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0003  row: A  column: 19
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 40.
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0003A19"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match             61.0%;  Score 12.2;  DB 28;  Length 40;
  Best Local Similarity   82.4%;  Pred. No. 1.9e+05;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTG 17
              ||||| || ||| ||||
Db         32 AGTAAAATGTATCTTTG 16


Search completed: September 23, 2004, 16:43:34
Job time : 1356 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 23, 2004, 13:59:29 ; Search time 1263 Seconds
(without alignments)
686.350 Million cell updates/sec

Title: US-10-798-923A-36
Perfect score: 20
Sequence: 1 agtaacatctatgtttggtt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 1774092

Minimum DB seq length: 0
Maximum DB seq length: 80

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	13.8	69.0	20	6	AR100077	AR100077 Sequence
c	2	13.8	69.0	29	8	AJ591807	AJ591807 Arabidops
	3	13.8	69.0	33	6	BD233882	BD233882 Novel met
	4	13.8	69.0	33	6	AX025318	AX025318 Sequence
	5	13.8	69.0	33	6	AX113479	AX113479 Sequence
	6	13.8	69.0	33	6	AX113614	AX113614 Sequence
	7	13.8	69.0	33	6	AX816281	AX816281 Sequence
	8	13.8	69.0	43	6	BD241774	BD241774 Nucleic a
c	9	13.6	68.0	28	6	AX074559	AX074559 Sequence
c	10	13.6	68.0	73	6	AX356671	AX356671 Sequence
c	11	13.6	68.0	73	6	BD175616	BD175616 Expressio
	12	13.2	66.0	24	6	AR049025	AR049025 Sequence
c	13	13.2	66.0	52	8	ATH521905	AJ521905 Arabidops
	14	13.2	66.0	65	6	AX486274	AX486274 Sequence
	15	12.8	64.0	30	6	AX180638	AX180638 Sequence
	16	12.8	64.0	34	6	AX136051	AX136051 Sequence
	17	12.8	64.0	34	6	BD014871	BD014871 Catalyst
c	18	12.8	64.0	43	6	AX484412	AX484412 Sequence
c	19	12.8	64.0	47	6	AR290152	AR290152 Sequence
c	20	12.8	64.0	51	6	AX165280	AX165280 Sequence
c	21	12.8	64.0	58	6	AX008365	AX008365 Sequence
c	22	12.8	64.0	58	6	BD218258	BD218258 Newcastle
c	23	12.8	64.0	74	6	AR147539	AR147539 Sequence
c	24	12.6	63.0	24	6	AR142867	AR142867 Sequence
c	25	12.6	63.0	24	6	AR194187	AR194187 Sequence
c	26	12.6	63.0	27	6	AR103699	AR103699 Sequence
c	27	12.6	63.0	27	6	BD129929	BD129929 Asthma-as
	28	12.6	63.0	34	6	I33654	I33654 Sequence 3
	29	12.6	63.0	35	6	A35752	A35752 Synthetic o
	30	12.6	63.0	35	6	AR169000	AR169000 Sequence
c	31	12.6	63.0	36	6	A35751	A35751 Synthetic o
c	32	12.6	63.0	36	6	AR168999	AR168999 Sequence
c	33	12.6	63.0	37	6	AR184398	AR184398 Sequence
c	34	12.6	63.0	50	6	AX453001	AX453001 Sequence
	35	12.6	63.0	55	6	AX485769	AX485769 Sequence
c	36	12.6	63.0	59	11	AF424885	AF424885 Mayetiola
	37	12.6	63.0	60	6	BD224826	BD224826 Novel pla
c	38	12.6	63.0	60	11	BV079655	BV079655 2712 Hess
c	39	12.6	63.0	63	6	AX918446	AX918446 Sequence
c	40	12.6	63.0	63	6	BD053979	BD053979 Sequence
	41	12.6	63.0	79	9	S47006	S47006 D3S745 {VNT
	42	12.6	63.0	80	6	AX241123	AX241123 Sequence
c	43	12.4	62.0	20	6	AX462717	AX462717 Sequence
	44	12.4	62.0	37	8	ATH531902	AJ531902 Arabidops
	45	12.4	62.0	48	1	SAMUPIRES	X59477 S.aureus pl
c	46	12.4	62.0	51	6	AX165120	AX165120 Sequence
	47	12.2	61.0	20	6	E10288	E10288 Primer for
	48	12.2	61.0	23	6	AX546452	AX546452 Sequence
	49	12.2	61.0	23	6	AX557293	AX557293 Sequence
	50	12.2	61.0	23	6	AX557377	AX557377 Sequence
	51	12.2	61.0	23	6	AX557402	AX557402 Sequence
	52	12.2	61.0	23	6	AX591113	AX591113 Sequence
	53	12.2	61.0	23	6	AX592503	AX592503 Sequence
	54	12.2	61.0	23	6	AX593006	AX593006 Sequence
	55	12.2	61.0	23	6	AX593146	AX593146 Sequence
	56	12.2	61.0	23	6	AX593481	AX593481 Sequence
	57	12.2	61.0	23	6	AX597476	AX597476 Sequence
	58	12.2	61.0	23	6	AX601686	AX601686 Sequence
	59	12.2	61.0	23	6	AX616987	AX616987 Sequence
	60	12.2	61.0	23	6	AX643861	AX643861 Sequence
	61	12.2	61.0	23	6	AX696027	AX696027 Sequence
	62	12.2	61.0	23	6	AX773019	AX773019 Sequence
	63	12.2	61.0	23	6	AX781403	AX781403 Sequence
	64	12.2	61.0	23	6	AX794420	AX794420 Sequence
	65	12.2	61.0	23	6	AX815459	AX815459 Sequence

	66	12.2	61.0	24	6	BD272920 BD272920 Methods.
	67	12.2	61.0	24	6	AX026175 AX026175 Sequence
	68	12.2	61.0	25	8	AJ587782 AJ587782 Arabidops
	69	12.2	61.0	26	6	E25013 E25013 Oligonucleo
	70	12.2	61.0	32	6	BD178885 BD178885 NR10 spli
c	71	12.2	61.0	34	6	I32311 I32311 Sequence 3
	72	12.2	61.0	38	6	AX203460 AX203460 Sequence
	73	12.2	61.0	42	6	AX225224 AX225224 Sequence
c	74	12.2	61.0	44	6	AX225223 AX225223 Sequence
	75	12.2	61.0	51	6	AX116449 AX116449 Sequence
	76	12.2	61.0	51	6	AX157286 AX157286 Sequence
	77	12.2	61.0	62	1	AF195101 AF195101 Sulfolobu
	78	12.2	61.0	69	8	AJ597178 AJ597178 Arabidops
c	79	12.2	61.0	69	8	ATH529042 AJ529042 Arabidops
c	80	12.2	61.0	69	8	ATH530578 AJ530578 Arabidops
	81	12.2	61.0	74	1	BAJ10130 AJ010130 Bacillus
	82	12.2	61.0	75	6	BD268125 BD268125 Stabilize
c	83	12.2	61.0	77	10	MUSJHCD3 M64559 Mouse JH-Cd
c	84	12.2	61.0	78	6	AX899169 AX899169 Sequence
c	85	12.2	61.0	78	6	BD034702 BD034702 Sequence
c	86	12	60.0	21	6	AX699149 AX699149 Sequence
	87	12	60.0	22	6	AR201837 AR201837 Sequence
c	88	12	60.0	24	6	AX164396 AX164396 Sequence
	89	12	60.0	30	6	AR028211 AR028211 Sequence
	90	12	60.0	30	6	AR138614 AR138614 Sequence
c	91	12	60.0	35	6	AX008525 AX008525 Sequence
c	92	12	60.0	35	6	BD218407 BD218407 Newcastle
	93	12	60.0	42	6	BD245088 BD245088 Regions o
	94	12	60.0	42	6	BD245098 BD245098 Regions o
	95	12	60.0	45	6	I55027 I55027 Sequence 51
	96	12	60.0	46	8	ATH520780 AJ520780 Arabidops
	97	12	60.0	47	6	AR290977 AR290977 Sequence
	98	12	60.0	52	6	AX058573 AX058573 Sequence
c	99	12	60.0	58	6	AR358864 AR358864 Sequence
c	100	12	60.0	59	8	AJ595268 AJ595268 Arabidops
c	101	12	60.0	72	9	AF043755 AF043755 Homo sapi
c	102	12	60.0	72	9	AF043761 AF043761 Homo sapi
c	103	12	60.0	72	9	AF043762 AF043762 Homo sapi
c	104	12	60.0	72	9	AF043798 AF043798 Homo sapi
c	105	12	60.0	72	9	AF043802 AF043802 Homo sapi
c	106	12	60.0	72	11	BX465922 BX465922 Arabidops
c	107	12	60.0	75	9	AF043756 AF043756 Homo sapi
c	108	12	60.0	75	9	AF043757 AF043757 Homo sapi
c	109	12	60.0	75	9	AF043758 AF043758 Homo sapi
c	110	12	60.0	75	9	AF043759 AF043759 Homo sapi
c	111	12	60.0	75	9	AF043760 AF043760 Homo sapi
c	112	12	60.0	75	9	AF043799 AF043799 Homo sapi
c	113	12	60.0	75	9	AF043800 AF043800 Homo sapi
c	114	12	60.0	75	9	AF043836 AF043836 Homo sapi
c	115	12	60.0	75	9	AF043837 AF043837 Homo sapi
c	116	12	60.0	75	9	AF043989 AF043989 Homo sapi
c	117	12	60.0	75	9	AF043991 AF043991 Homo sapi
c	118	12	60.0	75	9	AF043993 AF043993 Homo sapi
c	119	12	60.0	75	9	AF043995 AF043995 Homo sapi
c	120	12	60.0	78	9	AF043763 AF043763 Homo sapi
c	121	12	60.0	78	9	AF043834 AF043834 Homo sapi
c	122	12	60.0	78	9	AF043835 AF043835 Homo sapi
c	123	12	60.0	78	9	AF043838 AF043838 Homo sapi
c	124	12	60.0	78	9	AF043983 AF043983 Homo sapi
c	125	12	60.0	78	9	AF043985 AF043985 Homo sapi
c	126	12	60.0	78	9	AF043987 AF043987 Homo sapi
c	127	11.8	59.0	15	6	AX008487 AX008487 Sequence
c	128	11.8	59.0	15	6	BD218375 BD218375 Newcastle
	129	11.8	59.0	16	6	AR436140 AR436140 Sequence
c	130	11.8	59.0	17	6	AX226767 AX226767 Sequence
c	131	11.8	59.0	17	6	AX227579 AX227579 Sequence
c	132	11.8	59.0	17	6	AX227747 AX227747 Sequence
	133	11.8	59.0	18	6	I57056 I57056 Sequence 57
	134	11.8	59.0	18	6	AX105790 AX105790 Sequence
c	135	11.8	59.0	20	6	AR116568 AR116568 Sequence
c	136	11.8	59.0	20	6	BD074725 BD074725 Antisence
	137	11.8	59.0	25	12	AB086568 AB086568 Synthetic
c	138	11.8	59.0	26	6	AX814025 AX814025 Sequence
c	139	11.8	59.0	27	6	AX008357 AX008357 Sequence
c	140	11.8	59.0	27	6	BD218250 BD218250 Newcastle
	141	11.8	59.0	29	6	I80874 I80874 Sequence 10
	142	11.8	59.0	30	6	AX792894 AX792894 Sequence
	143	11.8	59.0	31	6	BD002365 BD002365 Gene comp
c	144	11.8	59.0	32	6	E14590 E14590 Oligonucleo
	145	11.8	59.0	32	6	I26668 I26668 Sequence 1
c	146	11.8	59.0	33	6	I26669 I26669 Sequence 2
	147	11.8	59.0	41	6	AX514674 AX514674 Sequence
	148	11.8	59.0	41	6	AX520614 AX520614 Sequence
c	149	11.8	59.0	42	6	AX060326 AX060326 Sequence
c	150	11.8	59.0	42	6	AX503648 AX503648 Sequence

ALIGNMENTS

```
RESULT 1
AR100077
LOCUS       AR100077                  20 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 34 from patent US 6080546.
ACCESSION   AR100077
VERSION     AR100077.1  GI:12810525
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Monia,B.P., Gaarde,W. and Cowsert,L.M.
  TITLE     Antisense modulation of MEKK5 expression
  JOURNAL   Patent: US 6080546-A 34 27-JUN-2000;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           69.0%;  Score 13.8;  DB 6;  Length 20;
  Best Local Similarity   88.2%;  Pred. No. 8.1e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTG 17
              ||||||||||  | ||||
Db          2 AGTAACATCTGTCTTTG 18


RESULT 2
AJ591807/c
LOCUS       AJ591807                  29 bp    DNA     linear   PLN 23-OCT-2003
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence, left border, clone
            592G08.
ACCESSION   AJ591807
VERSION     AJ591807.1  GI:37941431
KEYWORDS    left border; T-DNA flanking sequence.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Brunaud,V., Balzergue,S., Dubreucq,B., Aubourg,S., Samson,F.,
            Chauvin,S., Bechtold,N., Cruaud,C., DeRose,R., Pelletier,G.,
            Lepiniec,L., Caboche,M. and Lecharny,A.
  TITLE     T-DNA integration into the Arabidopsis genome depends on sequences
            of pre-insertion sites
  JOURNAL   EMBO Rep. 3 (12), 1152-1157 (2002)
  MEDLINE   22363535
   PUBMED   12446565
REFERENCE   2  (bases 1 to 29)
  AUTHORS   Balzergue,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-OCT-2003) Balzergue S., UMRGV, INRA/CNRS, 2 rue
            Gaston Cremieux, 91057 Evry cedex, FRANCE
```

```
COMMENT     PCR was performed on DNA from transformants of Arabidopsis thaliana
            plants from INRA (Versailles). The DNA fragment(s) resulting from
            the PCR were directly sequenced from the left or the right border
            to determine the genomic sequence flanking the insertion. T-DNA
            derived sequences were removed. Information to order the
            corresponding mutant line and a link to a database providing a
            graphical display of the insertion site are available at
            http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
            been generated in the framework of the French plant genomics
            program 'Genoplante' (http://www.genoplante.com and
            http://genoplante-info.infobiogen.fr).
FEATURES                Location/Qualifiers
     source             1. .29
                        /organism="Arabidopsis thaliana"
                        /mol_type="genomic DNA"
                        /cultivar="Wassillewskija"
                        /db_xref="taxon:3702"
                        /clone="592G08"
                        /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
     misc_feature       1. .29
                        /note="T-DNA flanking sequence
                        left border"
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 8;  Length 29;
  Best Local Similarity   88.2%;  Pred. No. 7.6e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 AACATCTATGTTTGGTT 20
              |||||| || |||||||
Db         27 AACATCAATTTTTGGTT 11


RESULT 3
BD233882
LOCUS       BD233882                 33 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Novel method for detecting acid-resistant microorganisms in feces.
ACCESSION   BD233882
VERSION     BD233882.1  GI:33043652
KEYWORDS    JP 2002529705-A/22.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 33)
  AUTHORS   Reiter,C., Cullmann,G., Friedrichs,U., Heppner,P., Lakner,M. and
            Ringeis,A.
  TITLE     Novel method for detecting acid-resistant microorganisms in feces
  JOURNAL   Patent: JP 2002529705-A 22 10-SEP-2002;
            CONNEX GMBH
COMMENT     OS   Artificial Sequence
            PN   JP 2002529705-A/22
            PD   10-SEP-2002
            PF   29-OCT-1999 JP 2000580001
            PR   29-OCT-1998 EP   98120517.2,06-NOV-1998 EP   98120687.3 PI
            CHRISTIAN REITER,GERHARD CULLMANN,ULRIKE FRIEDRICHS,PETRA  PI
            HEPPNER,
            PI   MERET LAKNER,ACHIM RINGEIS
            PC   G01N33/569,C07K16/12,G01N33/543,G01N33/577//C12P21/08,G01N33/
            PC   48,
            PC   (C12P21/08,C12R1:91)
            CC   Description of Artificial Sequence: Artificial Sequence FH
            Key             Location/Qualifiers
            FT   source          1. .33
            FT                   /organism='Artificial Sequence'.
FEATURES                Location/Qualifiers
     source             1. .33
                        /organism="synthetic construct"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:32630"
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 33;
  Best Local Similarity   88.2%;  Pred. No. 7.5e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 AACATCTATGTTTGGTT 20
              |||||  ||||||||||
Db         13 AACATTAATGTTTGGTT 29


RESULT 4
AX025318
LOCUS       AX025318                 33 bp    DNA     linear   PAT 16-SEP-2000
DEFINITION  Sequence 46 from Patent WO0026671.
ACCESSION   AX025318
VERSION     AX025318.1  GI:10187008
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Friedrichs,U., Heppner,P., Ringeis,A., Lakner,M., Cullmann,G. and
            Reiter,C.
  TITLE     Detection of acid-resistant micro-organisms in a stool
  JOURNAL   Patent: WO 0026671-A 46 11-MAY-2000;
            FRIEDRICHS ULRIKE (DE) ; CONNEX GMBH (DE) ; HEPPNER PETRA (DE) ;
            RINGEIS ACHIM (DE) ; LAKNER MERET (DE) ; CULLMANN GERHARD (DE) ;
            REITER CHRISTIAN (DE)
FEATURES                Location/Qualifiers
     source             1. .33
                        /organism="synthetic construct"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:32630"
                        /note="kunstliche Sequenz"
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 33;
  Best Local Similarity   88.2%;  Pred. No. 7.5e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 AACATCTATGTTTGGTT 20
              |||||  ||||||||||
Db         13 AACATTAATGTTTGGTT 29


RESULT 5
AX113479
LOCUS       AX113479                 33 bp    DNA     linear   PAT 01-MAY-2001
DEFINITION  Sequence 54 from Patent WO0127612.
ACCESSION   AX113479
VERSION     AX113479.1  GI:13939723
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Reiter,C., Cullmann,G., Lakner,M., Truee,A., Dehnert,S. and
            Schwartz,G.
  TITLE     Immuno-chromatographic rapid assay in order to detect
            acid-resistant microorganisms in the stool
  JOURNAL   Patent: WO 0127612-A 54 19-APR-2001;
            Connex Gesellschaft zur Optimierung von Forschung und Entwicklung
            mbH (DE)
FEATURES                Location/Qualifiers
     source             1. .33
                        /organism="synthetic construct"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:32630"
                        /note="CDR"
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 33;
  Best Local Similarity   88.2%;  Pred. No. 7.5e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 AACATCTATGTTTGGTT 20
```

```
              ||||| ||||||||||
Db         13 AACATTAATGTTTGGTT 29


RESULT 6
AX113614
LOCUS       AX113614                  33 bp    DNA     linear   PAT 01-MAY-2001
DEFINITION  Sequence 54 from Patent WO0127613.
ACCESSION   AX113614
VERSION     AX113614.1  GI:13939794
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Reiter,C., Cullmann,G., Heppner,P., Ringeis,A., Mueller,H. and
            Haindl,E.
  TITLE     Improved method for the detection of acid resistant microorganisms
            in a stool
  JOURNAL   Patent: WO 0127613-A 54 19-APR-2001;
            Connex Gesellschaft zur Optimierung von Forschung und Entwicklung
            (DE)
FEATURES             Location/Qualifiers
     source          1. .33
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="CDR"
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 33;
  Best Local Similarity   88.2%;  Pred. No. 7.5e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 AACATCTATGTTTGGTT 20
              ||||| ||||||||||
Db         13 AACATTAATGTTTGGTT 29


RESULT 7
AX816281
LOCUS       AX816281                  33 bp    DNA     linear   PAT 09-DEC-2003
DEFINITION  Sequence 54 from Patent EP1336850.
ACCESSION   AX816281
VERSION     AX816281.1  GI:39646788
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Reiter,C., Cullmann,G., Mueller,H., Heppner,P., Haindl,E. and
            Ringeis,A.
  TITLE     Improved method for the detection of acid resistant microorganisms
            in a stool
  JOURNAL   Patent: EP 1336850-A 54 20-AUG-2003;
            Connex Gesellschaft zur Optimierung von Forschung und Ent wicklung
            (DE)
FEATURES             Location/Qualifiers
     source          1. .33
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 33;
  Best Local Similarity   88.2%;  Pred. No. 7.5e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 AACATCTATGTTTGGTT 20
              ||||| ||||||||||
Db         13 AACATTAATGTTTGGTT 29


RESULT 8
BD241774
LOCUS       BD241774                  43 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Nucleic acids encoding taxus geranylgeranyl diphosphate synthase,
            and methods of use.
ACCESSION   BD241774
VERSION     BD241774.1  GI:33051544
KEYWORDS    JP 2002529077-A/8.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 43)
  AUTHORS   Croteau,R.B. and Hefner,J.L.
  TITLE     Nucleic acids encoding taxus geranylgeranyl diphosphate synthase,
            and methods of use
  JOURNAL   Patent: JP 2002529077-A 8 10-SEP-2002;
            WASHINGTON STATE UNIVERSITY RESEARCH FOUNDATION
COMMENT     OS   Artificial Sequence
            PN   JP 2002529077-A/8
            PD   10-SEP-2002
            PF   27-OCT-1999 JP 2000581172
            PR   05-NOV-1998 US    09/187050
            PI   RODNEY B CROTEAU,JERRY L HEFNER
            PC   C12N15/09,C12N5/10,C12N9/10//(C12N9/10,C12R1:91),C12N15/00, PC
              C12N5/00
            CC   Description of Artificial Sequence:PCR primer CC   PCR primer
            for synthesizing Tr295 truncation product FH   Key
            Location/Qualifiers
            FT   misc_difference (1). .(43).
FEATURES             Location/Qualifiers
     source          1. .43
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 43;
  Best Local Similarity   88.2%;  Pred. No. 7.2e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 AACATCTATGTTTGGTT 20
              || |||||||||| ||
Db          2 AAGATCTATGTTTGATT 18


RESULT 9
AX074559/c
LOCUS       AX074559                  28 bp    DNA     linear   PAT 06-FEB-2001
DEFINITION  Sequence 3 from Patent WO0104324.
ACCESSION   AX074559
VERSION     AX074559.1  GI:12710662
KEYWORDS    .
SOURCE      Clostridium butyricum
  ORGANISM  Clostridium butyricum
            Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
            Clostridium.
REFERENCE   1
  AUTHORS   Sarcabal,P., Croux,C. and Soucaille,P.
  TITLE     Method for preparing 1,3-propanediol by a recombinant
            micro-organism in the absence of coenzyme b12 or one of its
            precursors
  JOURNAL   Patent: WO 0104324-A 3 18-JAN-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Institut National des Sciences Appliquees de Toulouse (FR) ; Centre
            National De La Recherche Scientifique (FR)
FEATURES             Location/Qualifiers
     source          1. .28
                     /organism="Clostridium butyricum"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:1492"
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 28;
```

```
  Best Local Similarity   80.0%;  Pred. No. 9.5e+04;
  Matches   16;  Conservative   0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGTT 20
              | |||||| | |||||| ||
Db         26 AATAACATTTTTGTTTGTTT 7


RESULT 10
AX356671/c
LOCUS       AX356671                  73 bp    DNA     linear   PAT 13-FEB-2002
DEFINITION  Sequence 27 from Patent EP1176205.
ACCESSION   AX356671
VERSION     AX356671.1  GI:18674011
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Mueller,R., Thalhofer,J.P., Geipel,F., Hoelke,W., Glaser,S.,
            Eckstein,H., Kirschbaum,T. and Bommarius,B.
  TITLE     Expression of alkaline phosphatase in yeast
  JOURNAL   Patent: EP 1176205-A 27 30-JAN-2002;
            Roche Diagnostics GmbH (DE) ; F. HOFFMANN-LA ROCHE AG (CH)
FEATURES             Location/Qualifiers
     source          1. .73
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Artificial"
ORIGIN

  Query Match            68.0%;  Score 13.6;  DB 6;  Length 73;
  Best Local Similarity   80.0%;  Pred. No. 8.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGTT 20
              ||  || ||||| |||||||
Db         23 AGGTACTTCTATTTTTGGTT 4


RESULT 11
BD175616/c
LOCUS       BD175616                  73 bp    DNA     linear   PAT 18-MAR-2003
DEFINITION  Expression of alkaline phosphatase in yeast.
ACCESSION   BD175616
VERSION     BD175616.1  GI:29121314
KEYWORDS    JP 2002253269-A/26.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   Mueller,R., Thalhofer,J.P., Geipel,F., Hoelke,W., Glaser,S.,
            Eckstein,H., Kirschbaum,T. and Riebel,B.B.N.
  TITLE     Expression of alkaline phosphatase in yeast
  JOURNAL   Patent: JP 2002253269-A 26 10-SEP-2002;
            F HOFFMANN LA ROCHE AG
COMMENT     OS   Artificial Sequence
            PN   JP 2002253269-A/26
            PD   10-SEP-2002
            PF   23-JUL-2001 JP 2001222153
            PR   25-JUL-2000 DE   10036491.8
            PI   RAINER MUELLER,JOHANN PETER THALHOFER,FRANK GEIPEL,WERNER  PI
             HOELKE,
            PI   STEPHAN GLASER,HELLMUT ECKSTEIN,THOMAS KIRSCHBAUM, PI
            BETTINA BOMMARIUS NEE RIEBEL
            PC   C12N15/09,C12N1/19,C12N9/16//(C12N1/19,C12R1:645),(C12N1/19,
            PC   C12R1:78),
            PC   (C12N9/16,C12R1:645),(C12N9/16,C12R1:78),C12N15/00 CC
            Description of Artificial Sequence: Artificial FH   Key
            Location/Qualifiers
            FT   source          1. .73
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .73
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match            68.0%;  Score 13.6;  DB 6;  Length 73;
  Best Local Similarity   80.0%;  Pred. No. 8.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGTT 20
              ||  || ||||| |||||||
Db         23 AGGTACTTCTATTTTTGGTT 4


RESULT 12
AR049025
LOCUS       AR049025                  24 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 14 from patent US 5824501.
ACCESSION   AR049025
VERSION     AR049025.1  GI:6005064
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Ellis,N.A., German,J. and Groden,J.
  TITLE     Nucleic acids of the bloom's syndrome gene
  JOURNAL   Patent: US 5824501-A 14 20-OCT-1998;
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match            66.0%;  Score 13.2;  DB 6;  Length 24;
  Best Local Similarity   83.3%;  Pred. No. 1.5e+05;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGG 18
              |||| |||| ||| ||||
Db          6 AGTACCATCAATGATTGG 23


RESULT 13
ATH521905/c
LOCUS       ATH521905                 52 bp    DNA     linear   PLN 29-MAR-2003
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence, left border, clone
            280G09.
ACCESSION   AJ521905
VERSION     AJ521905.1  GI:26790141
KEYWORDS    left border; T-DNA flanking sequence.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Brunaud,V., Balzergue,S., Dubreucq,B., Aubourg,S., Samson,F.,
            Chauvin,S., Bechtold,N., Cruaud,C., DeRose,R., Pelletier,G.,
            Lepiniec,L., Caboche,M. and Lecharny,A.
  TITLE     T-DNA integration into the Arabidopsis genome depends on sequences
            of pre-insertion sites
  JOURNAL   EMBO Rep. 3 (12), 1152-1157 (2002)
  MEDLINE   22363535
   PUBMED   12446565
REFERENCE   2  (bases 1 to 52)
  AUTHORS   Balzergue,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue
            Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT     PCR was performed on DNA from transformants of Arabidopsis thaliana
```

```
            plants from INRA (Versailles). The DNA fragment(s) resulting from
            the PCR were directly sequenced from the left or the right border
            to determine the genomic sequence flanking the insertion. T-DNA
            derived sequences were removed. Information to order the
            corresponding mutant line and a link to a database providing a
            graphical display of the insertion site are available at
            http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
            been generated in the framework of the French plant genomics
            program 'Genoplante' (http://www.genoplante.com and
            http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /cultivar="Wassillewskija"
                     /db_xref="taxon:3702"
                     /clone="280G09"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
     misc_feature    1. .52
                     /note="T-DNA flanking sequence
                     left border"
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 8;  Length 52;
  Best Local Similarity   83.3%;  Pred. No. 1.3e+05;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 TAACATCTATGTTTGGTT 20
              ||| |||||| |||| ||
Db         18 TAATATCTATTTTTGATT 1


RESULT 14
AX486274
LOCUS       AX486274                  65 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 3574 from Patent WO02053728.
ACCESSION   AX486274
VERSION     AX486274.1  GI:22320490
KEYWORDS    .
SOURCE      Candida albicans
  ORGANISM  Candida albicans
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE   1
  AUTHORS   Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
  TITLE     Gene disruption methodologies for drug target discovery
  JOURNAL   Patent: WO 02053728-A 3574 11-JUL-2002;
            Elitra Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Candida albicans"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:5476"
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 65;
  Best Local Similarity   83.3%;  Pred. No. 1.3e+05;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGT 19
              |||||||  | |||||||
Db          7 GTAACATTCAAGTTTGGT 24


RESULT 15
AX180638
LOCUS       AX180638                  30 bp    DNA     linear   PAT 06-AUG-2001
DEFINITION  Sequence 216 from Patent WO0146391.
ACCESSION   AX180638
VERSION     AX180638.1  GI:15132524
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Osbourn,A.E., Haralampidis,K. and Bryan,G.T.
  TITLE     Plant gene
  JOURNAL   Patent: WO 0146391-A 216 28-JUN-2001;
            Plant Bioscience Limited (GB)
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Primer"
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 30;
  Best Local Similarity   87.5%;  Pred. No. 2.2e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGGTT 20
              ||||| |||||||  ||
Db          7 ACATCCATGTTTGTTT 22


RESULT 16
AX136051
LOCUS       AX136051                  34 bp    DNA     linear   PAT 30-MAY-2001
DEFINITION  Sequence 3 from Patent EP1096014.
ACCESSION   AX136051
VERSION     AX136051.1  GI:14272474
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Chen,P., Kan,C.C., Luo,C., Margosiak,S., O'Connor,P.,
            Tempczyk-Russel,A., Nguyen,B., Sarup,J.C., Gaur,S., Anderson,M.B.,
            Deng,Y.L., Lundgren,K. and Register,J.
  TITLE     Catalytic domain of the human effector cell cycle checkpoint
            protein kinase, chk1, materials and methods for identification of
            inhibitors thereof
  JOURNAL   Patent: EP 1096014-A 3 02-MAY-2001;
            Agouron Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .34
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="PCR primer"
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 34;
  Best Local Similarity   87.5%;  Pred. No. 2.2e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTT 16
              |||| ||||||| |||
Db          7 AGTACCATCTATCTTT 22


RESULT 17
BD014871
LOCUS       BD014871                  34 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Catalyst domain of human effector cell cycle check point protein
            kinase Chk1, and substance for identifying the inhibitor and
            identification method.
ACCESSION   BD014871
VERSION     BD014871.1  GI:22555678
KEYWORDS    JP 2001161387-A/2.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 34)
  AUTHORS   Chen,P., Kan,C.C., Luo,C., Margosiak,S., OGconnor,P., Russell,A.T.,
```

```
            Nguyen,B., Sarup,J.C., Gaur,S., Anderson,M.B., Deng,Y.L.,
            Lundgren,K. and Register,J.
  TITLE     Catalyst domain of human effector cell cycle check point protein
            kinase Chk1, and substance for identifying the inhibitor and
            identification method
  JOURNAL   Patent: JP 2001161387-A 2 19-JUN-2001;
            AGOURON PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2001161387-A/2
            PD   19-JUN-2001
            PF   01-NOV-2000 JP 2000335268
            PR   01-NOV-1999 US    60/162887,14-DEC-1999 US    09/460421 PI
            PING CHEN,CHEN CHEN KAN,CHUN LUO,STEVEN MARGOSIAK,PATRICK  PI
            O'CONNOR,
            PI   ANNA TEMPCZYK RUSSELL,BINH NGUYEN,JAY CHAND SARUP,SMITA GAUR,
            PI   MARK BRIAN ANDERSON,YA LI DENG,KAREN LUNDGREN,JAMES REGISTER
            PC   C12N15/09,C07K19/00,C12N1/21,C12N5/10,C12N7/00,C12N9/12,C12N9/
            PC   99,
            PC   C12Q1/48//(C12N1/21,C12R1:19),(C12N5/10,C12R1:91),C12N15/00,
            PC   C12N5/00,
            PC   (C12N5/00,C12R1:91)
            CC   Description of Artificial Sequence: PCR primer FH   Key
                 Location/Qualifiers
            FT   source          1. .34
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .34
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match           64.0%;  Score 12.8;  DB 6;  Length 34;
  Best Local Similarity  87.5%;  Pred. No. 2.2e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTT 16
              |||| ||||||| |||
Db          7 AGTACCATCTATCTTT 22


RESULT 18
AX484412/c
LOCUS       AX484412                  43 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 1712 from Patent WO02053728.
ACCESSION   AX484412
VERSION     AX484412.1  GI:22318764
KEYWORDS    .
SOURCE      Candida albicans
  ORGANISM  Candida albicans
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE   1
  AUTHORS   Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
  TITLE     Gene disruption methodologies for drug target discovery
  JOURNAL   Patent: WO 02053728-A 1712 11-JUL-2002;
            Elitra Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .43
                     /organism="Candida albicans"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:5476"
ORIGIN

  Query Match           64.0%;  Score 12.8;  DB 6;  Length 43;
  Best Local Similarity  87.5%;  Pred. No. 2.1e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGGTT 20
              | |||||||||| ||
Db         31 AGATCTATGTTTGTTT 16


RESULT 19
AR290152/c
LOCUS       AR290152                  47 bp    DNA     linear   PAT 12-JUN-2003
DEFINITION  Sequence 1887 from patent US 6537751.
ACCESSION   AR290152
VERSION     AR290152.1  GI:31677436
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Cohen,D., Chumakov,I. and Blumenfeld,M.
  TITLE     Biallelic markers for use in constructing a high density
            disequilibrium map of the human genome
  JOURNAL   Patent: US 6537751-A 1887 25-MAR-2003;
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           64.0%;  Score 12.8;  DB 6;  Length 47;
  Best Local Similarity  87.5%;  Pred. No. 2.1e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGGTT 20
              |||| ||||||||  ||
Db         43 ACATTTATGTTTGTTT 28


RESULT 20
AX165280/c
LOCUS       AX165280                  51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 475 from Patent WO0138586.
ACCESSION   AX165280
VERSION     AX165280.1  GI:14546109
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0138586-A 475 31-MAY-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     variation       26
                     /note="single nucleotide polymorphism
                     Accession number cg44005525"
ORIGIN

  Query Match           64.0%;  Score 12.8;  DB 6;  Length 51;
  Best Local Similarity  87.5%;  Pred. No. 2.1e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 TAACATCTATGTTTGG 18
              |||||||||||  |||
Db         38 TAACATCTATGAGTGG 23


RESULT 21
AX008365/c
LOCUS       AX008365                  58 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 17 from Patent WO9966045.
ACCESSION   AX008365
VERSION     AX008365.1  GI:9995921
KEYWORDS    .
```

```
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Gielkens,A.L., Koch,G., De Leeuw,O. and Peeters,B.P.
  TITLE     Newcastle disease virus infectious clones, vaccines and diagnostic
            assays
  JOURNAL   Patent: WO 9966045-A 17 23-DEC-1999;
            GIELKENS ARNOUD LEONARD JOSEF (NL); KOCH GUUS (NL); LEEUW OLAV SVEN
            DE (NL); PEETERS BERNARDUS PETRUS HUBER (NL); STICHTING DIENST
            LANDBOUWKUNDI (NL)
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
     primer_bind     1. .58
                     /note="'Primer BGL5F2'
                     primer"
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 58;
  Best Local Similarity   87.5%;  Pred. No. 2e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGGTT 20
              | |||| |||||||||
Db         50 AAATCTTTGTTTGGTT 35


RESULT 22
BD218258/c
LOCUS       BD218258                  58 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Newcastle disease virus infectious clones, vaccines and diagnostic
            assays.
ACCESSION   BD218258
VERSION     BD218258.1  GI:33028028
KEYWORDS    JP 2002518012-A/17.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   Peeters,B.P.H., Leeuw,O.S.D., Koch,G. and Gielkens,A.L.J.
  TITLE     Newcastle disease virus infectious clones, vaccines and diagnostic
            assays
  JOURNAL   Patent: JP 2002518012-A 17 25-JUN-2002;
            ID LELYSTAD INSTITUUT VOOR DIERHOUDERIJ EN DIERGEZONDHEID BV
COMMENT     OS   Artificial Sequence
            PN   JP 2002518012-A/17
            PD   25-JUN-2002
            PF   17-JUN-1999 JP 2000554854
            PR   19-JUN-1998 EP   98202054.7
            PI   BERNARDUS PETRUS HUBERTUS PEETERS,OLAV SVEN
            DE LEEUW,GUUS KOCH,
            PI   ARNOUD LEONARD JOSEF GIELKENS
            PC   C12N15/09,A61K39/17,A61K48/00,A61P31/12,C12N7/00,C12Q1/70, PC
             C12N15/00
            CC   /note='Primer BGL5F2'
            CC   Description of Artificial Sequence: primer
            FH   Key             Location/Qualifiers
            FT   primer_bind     (1). .(58).
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 58;
  Best Local Similarity   87.5%;  Pred. No. 2e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGGTT 20
              | |||| |||||||||
Db         50 AAATCTTTGTTTGGTT 35


RESULT 23
AR147539/c
LOCUS       AR147539                  74 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 53 from patent US 6221597.
ACCESSION   AR147539
VERSION     AR147539.1  GI:15111342
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 74)
  AUTHORS   Roberts,C.J.
  TITLE     Essential genes of yeast as targets for antifungal agents,
            herbicides, insecticides and anti-proliferative drugs
  JOURNAL   Patent: US 6221597-A 53 24-APR-2001;
FEATURES             Location/Qualifiers
     source          1. .74
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 74;
  Best Local Similarity   87.5%;  Pred. No. 2e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ACATCTATGTTTGGTT 20
              ||||| || |||||||
Db         23 ACATCCATCTTTGGTT 8


RESULT 24
AR142867/c
LOCUS       AR142867                  24 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 5 from patent US 6204019.
ACCESSION   AR142867
VERSION     AR142867.1  GI:15104153
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   O'Dwyer,K.M., Warren,R. and Perry,C.
  TITLE     Sec A2 from Streptococcus pneumoniae
  JOURNAL   Patent: US 6204019-A 5 20-MAR-2001;
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 6;  Length 24;
  Best Local Similarity   78.9%;  Pred. No. 2.9e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGTT 20
              |||||||||| ||  |||
Db         23 GTAACATCTAGTTTATGTT 5


RESULT 25
AR194187/c
LOCUS       AR194187                  24 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 5 from patent US 6348342.
ACCESSION   AR194187
VERSION     AR194187.1  GI:20240779
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
```

```
REFERENCE   1  (bases 1 to 24)
  AUTHORS   O'Dwyer,K.M., Warren,R. and Perry,C.
  TITLE     Compounds
  JOURNAL   Patent: US 6348342-A 5 19-FEB-2002;
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           63.0%;  Score 12.6;  DB 6;  Length 24;
  Best Local Similarity   78.9%;  Pred. No. 2.9e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGTT 20
              ||||||||||  ||  |||
Db         23 GTAACATCTAGTTTATGTT 5


RESULT 26
AR103699/c
LOCUS       AR103699                  27 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 223 from patent US 6087485.
ACCESSION   AR103699
VERSION     AR103699.1  GI:12815287
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Brooks-Wilson,A.R., Buckler,A., Cardon,L., Carey,A.H., Galvin,M.,
            Miller,A. and North,M.
  TITLE     Asthma related genes
  JOURNAL   Patent: US 6087485-A 223 11-JUL-2000;
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           63.0%;  Score 12.6;  DB 6;  Length 27;
  Best Local Similarity   78.9%;  Pred. No. 2.8e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              ||||||||||  |  ||||
Db         27 AGTAACATCTCAGCCTGGT 9


RESULT 27
BD129929/c
LOCUS       BD129929                  27 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Asthma-associated gene.
ACCESSION   BD129929
VERSION     BD129929.1  GI:23224874
KEYWORDS    JP 2002500895-A/219.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Wilson,A.R.B., Buckler,A., Cardon,L., Carey,A.H., Galvin,M.,
            Miller,A. and North,M.
  TITLE     Asthma-associated gene
  JOURNAL   Patent: JP 2002500895-A 219 15-JAN-2002;
            AXYS PHARMACEUTICALS INC
COMMENT     OS   Unidentified
            PN   JP 2002500895-A/219
            PD   15-JAN-2002
            PF   21-JAN-1998 JP 2000528715
            PI   ANGELA R BROOKS WILSON,ALAN BUCKLER,LON
            CARDON,ALISOUN H CAREY,
            PI   MARGARET GALVIN,ANDREW MILLER,MICHAEL NORTH
            PC   C12Q1/68,A01K67/027,C07K14/47,C12N15/09,C12N15/00 CC
            Strandedness: Single;
            CC   Topology: Linear;
            CC   Asthma-associated gene
            FH   Key             Location/Qualifiers
            FT   source          1. .27
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match           63.0%;  Score 12.6;  DB 6;  Length 27;
  Best Local Similarity   78.9%;  Pred. No. 2.8e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              ||||||||||  |  ||||
Db         27 AGTAACATCTCAGCCTGGT 9


RESULT 28
I33654
LOCUS       I33654                    34 bp    DNA     linear   PAT 06-FEB-1997
DEFINITION  Sequence 3 from patent US 5593857.
ACCESSION   I33654
VERSION     I33654.1  GI:1824445
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 34)
  AUTHORS   Higaki,J.N., Tischer,E.G., Cordell,B. and Thompson,S.A.
  TITLE     Production of homogeneous truncated CNTF
  JOURNAL   Patent: US 5593857-A 3 14-JAN-1997;
FEATURES             Location/Qualifiers
     source          1. .34
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           63.0%;  Score 12.6;  DB 6;  Length 34;
  Best Local Similarity   78.9%;  Pred. No. 2.7e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGTT 20
              |||  |  ||  ||||||  |||
Db         15 GTACCTTCCATGTTTTGTT 33


RESULT 29
A35752
LOCUS       A35752                    35 bp    DNA     linear   PAT 03-DEC-1996
DEFINITION  Synthetic oligo 50.
ACCESSION   A35752
VERSION     A35752.1  GI:1927123
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 35)
  AUTHORS   .
  TITLE     THYMIDINE PHOSPHORYLASE FOR USE IN THE MODULATION OF CELLULAR
            PROLIFERATION OR CHEMOTAXIS
  JOURNAL   Patent: WO 9308273-A 50 29-APR-1993;
FEATURES             Location/Qualifiers
     source          1. .35
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN
```

```
  Query Match             63.0%;  Score 12.6;  DB 6;  Length 35;
  Best Local Similarity   78.9%;  Pred. No. 2.7e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              | ||||||||  || ||| |
Db          5 AATAACATCTTTGCTTGTT 23


RESULT 30
AR169000
LOCUS       AR169000                  35 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 50 from patent US 6290953.
ACCESSION   AR169000
VERSION     AR169000.1  GI:17906699
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 35)
  AUTHORS   Ballance,D.J., Courtney,M.G., Finnis,C.J.A. and Sleep,D.
  TITLE     Modulation of cellular proliferation with thymidine phosphorylase
  JOURNAL   Patent: US 6290953-A 50 18-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .35
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 6;  Length 35;
  Best Local Similarity   78.9%;  Pred. No. 2.7e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              | ||||||||  || ||| |
Db          5 AATAACATCTTTGCTTGTT 23


RESULT 31
A35751/c
LOCUS       A35751                    36 bp    DNA     linear   PAT 03-DEC-1996
DEFINITION  Synthetic oligo 49.
ACCESSION   A35751
VERSION     A35751.1  GI:1927122
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 36)
  AUTHORS   .
  TITLE     THYMIDINE PHOSPHORYLASE FOR USE IN THE MODULATION OF CELLULAR
            PROLIFERATION OR CHEMOTAXIS
  JOURNAL   Patent: WO 9308273-A 49 29-APR-1993;
FEATURES             Location/Qualifiers
     source          1. .36
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 6;  Length 36;
  Best Local Similarity   78.9%;  Pred. No. 2.7e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              | ||||||||  || ||| |
Db         35 AATAACATCTTTGCTTGTT 17


RESULT 32
AR168999/c
LOCUS       AR168999                  36 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 49 from patent US 6290953.
ACCESSION   AR168999
VERSION     AR168999.1  GI:17906697
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 36)
  AUTHORS   Ballance,D.J., Courtney,M.G., Finnis,C.J.A. and Sleep,D.
  TITLE     Modulation of cellular proliferation with thymidine phosphorylase
  JOURNAL   Patent: US 6290953-A 49 18-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .36
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 6;  Length 36;
  Best Local Similarity   78.9%;  Pred. No. 2.7e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              | ||||||||  || ||| |
Db         35 AATAACATCTTTGCTTGTT 17


RESULT 33
AR184398/c
LOCUS       AR184398                  37 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 11 from patent US 6346378.
ACCESSION   AR184398
VERSION     AR184398.1  GI:20230363
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 37)
  AUTHORS   Stanley,C.John., Orum,H. and Jorgensen,M.
  TITLE     Nucleic acid analogs with a chelating functionality
  JOURNAL   Patent: US 6346378-A 11 12-FEB-2002;
FEATURES             Location/Qualifiers
     source          1. .37
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 6;  Length 37;
  Best Local Similarity   78.9%;  Pred. No. 2.7e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGTT 20
              || ||| |||| ||| |||
Db         32 GTCACAACTATTTTTAGTT 14


RESULT 34
AX453001/c
LOCUS       AX453001                  50 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 15 from Patent WO0244195.
ACCESSION   AX453001
VERSION     AX453001.1  GI:21712580
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Hayashizaki,Y.
  TITLE     Method for base sequencing and biologically active nucleic acids
  JOURNAL   Patent: WO 0244195-A 15 06-JUN-2002;
            RIKEN (JP)
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
```

```
                     /db_xref="taxon:32630"
                     /note="a spacer"
ORIGIN

  Query Match           63.0%;  Score 12.6;  DB 6;  Length 50;
  Best Local Similarity  78.9%;  Pred. No. 2.6e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              |||  |||||||||  |||
Db         21 AGTGTCATCTATGTCGGGT 3


RESULT 35
AX485769
LOCUS       AX485769                 55 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 3069 from Patent WO02053728.
ACCESSION   AX485769
VERSION     AX485769.1  GI:22319985
KEYWORDS    .
SOURCE      Candida albicans
  ORGANISM  Candida albicans
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE   1
  AUTHORS   Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
  TITLE     Gene disruption methodologies for drug target discovery
  JOURNAL   Patent: WO 02053728-A 3069 11-JUL-2002;
            Elitra Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Candida albicans"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:5476"
ORIGIN

  Query Match           63.0%;  Score 12.6;  DB 6;  Length 55;
  Best Local Similarity  78.9%;  Pred. No. 2.6e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGTT 20
              | ||| ||| ||||||| ||
Db         22 GAAACTTCTTTGTTTGTTT 40


RESULT 36
AF424885/c
LOCUS       AF424885                 59 bp    DNA     linear   STS 06-SEP-2002
DEFINITION  Mayetiola destructor chromosome X2 map between G15-1 and
            EAC/MCTA-201, sequence tagged site.
ACCESSION   AF424885
VERSION     AF424885.1  GI:16660403
KEYWORDS    STS.
SOURCE      Mayetiola destructor (Hessian fly)
  ORGANISM  Mayetiola destructor
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Nematocera; Sciaroidea;
            Cecidomyiidae; Mayetiola.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Rider,S.D. Jr., Sun,W., Ratcliffe,R.H. and Stuart,J.J.
  TITLE     Chromosome landing near avirulence gene vH13 in the Hessian fly
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 59)
  AUTHORS   Stuart,J.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-SEP-2001) Entomology, Purdue University, 1158 Smith
            Hall, West Lafayette, IN 47907-1158, USA
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="Mayetiola destructor"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:39758"
                     /chromosome="X2"
                     /map="between G15-1 and EAC/MCTA-201"
     STS             1. .59
                     /note="polymorphic AFLP marker EAC/MCTG-82"
ORIGIN

  Query Match           63.0%;  Score 12.6;  DB 11;  Length 59;
  Best Local Similarity  78.9%;  Pred. No. 2.5e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGTT 20
              | ||||| ||||| || ||
Db         45 GGAACATATATGTATGCTT 27


RESULT 37
BD224826
LOCUS       BD224826                 60 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Novel plant acyltransferases.
ACCESSION   BD224826
VERSION     BD224826.1  GI:33034596
KEYWORDS    JP 2002525105-A/175.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Lassner,M.W., Emig,R.A., Ruezinsky,D.M. and Eenennaam,A.V.
  TITLE     Novel plant acyltransferases
  JOURNAL   Patent: JP 2002525105-A 175 13-AUG-2002;
            CALGENE LLC
COMMENT     OS   Artificial Sequence
            PN   JP 2002525105-A/175
            PD   13-AUG-2002
            PF   24-SEP-1999 JP 2000572337
            PR   25-SEP-1998 US    60/101939
            PI   MICHAEL W LASSNER,ROBIN A EMIG,DIANE M RUEZINSKY,ALISON VAN
            PI   EENENNAAM
            PC   C12N15/09,A01H5/00,C12N5/10//C12N9/10,C12N15/00,C12N5/00 CC
            Description of Artificial Sequence:Synthetic Oligonucleotide FH
            Key             Location/Qualifiers
            FT   source          1. .60
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match           63.0%;  Score 12.6;  DB 6;  Length 60;
  Best Local Similarity  78.9%;  Pred. No. 2.5e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGTT 20
              | ||||| |||| | ||||
Db         22 GAAACATATATGGTCGGTT 40


RESULT 38
BV079655/c
LOCUS       BV079655                 60 bp    DNA     linear   STS 17-SEP-2003
DEFINITION  2712 Hessian fly genomic DNA Mayetiola destructor STS genomic,
            sequence tagged site.
ACCESSION   BV079655
VERSION     BV079655.1  GI:34787404
KEYWORDS    STS.
SOURCE      Mayetiola destructor (Hessian fly)
  ORGANISM  Mayetiola destructor
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Nematocera; Sciaroidea;
            Cecidomyiidae; Mayetiola.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Behura,S.K., Rider,S.D., Valicente,F.H. and Staurt,J.J.
  TITLE     Hessian fly STS markers
```

```
  JOURNAL   Unpublished (2003)
COMMENT
            Contact: Jeff Stuart
            Department of Entomology
            Purdue University
            901 W State St., West Lafayette, IN 47907, USA
            Primer A: None provided
            Primer B: None provided
            Protocol:
                        Template:    20-30ng
                        Primer:        each 20 pmoles
                        dNTPs:       each 200Um
                        Taq Pol:      0.1 units/ul
                        Total Vol:   25 ul

            Buffer:
                      MgCl2:    2.5mM
                      KCl:        50 mM
                      Tris-Cl:  10mM
                      pH:         8.3.
FEATURES              Location/Qualifiers
     source           1. .60
                      /organism="Mayetiola destructor"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:39758"
                      /clone_lib="Hessian fly genomic DNA"
     STS              <1. .>60
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 11;  Length 60;
  Best Local Similarity   78.9%;  Pred. No. 2.5e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGTT 20
              | ||||| ||||| || ||
Db         46 GGAACATATATGTATGCTT 28


RESULT 39
AX918446/c
LOCUS       AX918446                  63 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 34309 from Patent EP1033401.
ACCESSION   AX918446
VERSION     AX918446.1  GI:40212235
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
  TITLE     Expressed sequence tags and encoded human proteins
  JOURNAL   Patent: EP 1033401-A 34309 06-SEP-2000;
            Genset (FR)
FEATURES              Location/Qualifiers
     source           1. .63
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 6;  Length 63;
  Best Local Similarity   78.9%;  Pred. No. 2.5e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              || ||| || || ||||||
Db         43 AGCAACTTCCATTTTTGGT 25


RESULT 40
BD053979/c
LOCUS       BD053979                  63 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Sequence tag and encoded human protein.
ACCESSION   BD053979
VERSION     BD053979.1  GI:22599585
KEYWORDS    JP 2001269182-A/30225.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 63)
  AUTHORS   Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
  TITLE     Sequence tag and encoded human protein
  JOURNAL   Patent: JP 2001269182-A 30225 02-OCT-2001;
            GENSET
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001269182-A/30225
            PD   02-OCT-2001
            PF   24-FEB-2000 JP 2000118773
            PR   26-FEB-1999 US    60/122487
            PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
            PI   JORDAN
            PC   C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
             C12N5/10,
            PC   C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
              G06F15/40
            CC
            FH   Key             Location/Qualifiers.
FEATURES              Location/Qualifiers
     source           1. .63
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 6;  Length 63;
  Best Local Similarity   78.9%;  Pred. No. 2.5e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AGTAACATCTATGTTTGGT 19
              || ||| || || ||||||
Db         43 AGCAACTTCCATTTTTGGT 25


RESULT 41
S47006
LOCUS       S47006                    79 bp    DNA     linear   PRI 08-MAY-1993
DEFINITION  D3S745 {VNTR repeat element} [human, Genomic, 79 nt].
ACCESSION   S47006
VERSION     S47006.1  GI:258800
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   Latif,F., Modi,W.S., Duh,F.M., Schmidt,L., Li,H., Geil,L.,
            Orcutt,M.L., Heppell-Parton,A., Rabbitts,P.H., Linehan,W.M. et,al.
  TITLE     Molecular and genetic characterization and physical mapping of 11
            new markers detecting multiallele restriction fragment length
            polymorphisms on the short arm of human chromosome 3
  JOURNAL   Hum. Genet. 90 (1-2), 17-22 (1992)
  MEDLINE   93052228
   PUBMED   1358787
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 117048] from the original journal article.
            This sequence comes from Fig. 2.
            Map location: 3p26.
FEATURES              Location/Qualifiers
     source           1. .79
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
     gene             1. .79
                      /gene="D3S745"
```

```
ORIGIN

  Query Match            63.0%;  Score 12.6;  DB 9;  Length 79;
  Best Local Similarity  78.9%;  Pred. No. 2.4e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGTT 20
              |||| |  | |||||||||
Db          3 GTAAAAGTTGTGTTTGGTT 21


RESULT 42
AX241123
LOCUS       AX241123                  80 bp    DNA     linear   PAT 26-SEP-2001
DEFINITION  Sequence 361 from Patent WO0160975.
ACCESSION   AX241123
VERSION     AX241123.1  GI:15797998
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Roemer,T., Jiang,B., Boone,C. and Bussey,H.
  TITLE     Gene disruption methodologies for drug target discovery
  JOURNAL   Patent: WO 0160975-A 361 23-AUG-2001;
            Elitra Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .80
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="DNA primer"
ORIGIN

  Query Match            63.0%;  Score 12.6;  DB 6;  Length 80;
  Best Local Similarity  78.9%;  Pred. No. 2.4e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 GTAACATCTATGTTTGGTT 20
              | ||| ||| |||||| ||
Db         22 GAAACTTCTTTGTTTGTTT 40


RESULT 43
AX462717/c
LOCUS       AX462717                  20 bp    DNA     linear   PAT 15-JUL-2002
DEFINITION  Sequence 461 from Patent EP1217079.
ACCESSION   AX462717
VERSION     AX462717.1  GI:21885943
KEYWORDS    .
SOURCE      Aegilops tauschii
  ORGANISM  Aegilops tauschii
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Aegilops.
REFERENCE   1
  AUTHORS   Bernard,M., Sourdille,P. and Guyomarch,H.
  TITLE     Microsatellite markers from Triticum tauschii
  JOURNAL   Patent: EP 1217079-A 461 26-JUN-2002;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="Aegilops tauschii"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:37682"
ORIGIN

  Query Match            62.0%;  Score 12.4;  DB 6;  Length 20;
  Best Local Similarity  92.9%;  Pred. No. 3.7e+05;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 CATCTATGTTTGGT 19
              || ||||||||||
Db         20 CAGCTATGTTTGGT 7


RESULT 44
ATH531902
LOCUS       ATH531902                 37 bp    DNA     linear   PLN 29-MAR-2003
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence, left border, clone
            247D02.
ACCESSION   AJ531902
VERSION     AJ531902.1  GI:26800162
KEYWORDS    left border; T-DNA flanking sequence.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Brunaud,V., Balzergue,S., Dubreucq,B., Aubourg,S., Samson,F.,
            Chauvin,S., Bechtold,N., Cruaud,C., DeRose,R., Pelletier,G.,
            Lepiniec,L., Caboche,M. and Lecharny,A.
  TITLE     T-DNA integration into the Arabidopsis genome depends on sequences
            of pre-insertion sites
  JOURNAL   EMBO Rep. 3 (12), 1152-1157 (2002)
  MEDLINE   22363535
   PUBMED   12446565
REFERENCE   2  (bases 1 to 37)
  AUTHORS   Balzergue,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue
            Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT     PCR was performed on DNA from transformants of Arabidopsis thaliana
            plants from INRA (Versailles). The DNA fragment(s) resulting from
            the PCR were directly sequenced from the left or the right border
            to determine the genomic sequence flanking the insertion. T-DNA
            derived sequences were removed. Information to order the
            corresponding mutant line and a link to a database providing a
            graphical display of the insertion site are available at
            http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
            been generated in the framework of the French plant genomics
            program 'Genoplante' (http://www.genoplante.com and
            http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .37
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /cultivar="Wassillewskija"
                     /db_xref="taxon:3702"
                     /clone="247D02"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
     misc_feature    1. .37
                     /note="T-DNA flanking sequence
                     left border"
ORIGIN

  Query Match            62.0%;  Score 12.4;  DB 8;  Length 37;
  Best Local Similarity  92.9%;  Pred. No. 3.4e+05;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 ATCTATGTTTGGTT 20
              | ||||||||||||
Db          9 AACTATGTTTGGTT 22


RESULT 45
SAMUPIRES
LOCUS       SAMUPIRES                 48 bp    DNA     linear   BCT 02-MAR-1992
DEFINITION  S.aureus plasmid DNA for part of mupirocin resistance gene (XhoI
            site).
ACCESSION   X59477
VERSION     X59477.1  GI:46621
KEYWORDS    resistance gene.
SOURCE      Staphylococcus aureus
  ORGANISM  Staphylococcus aureus
```

```
           Bacteria; Firmicutes; Bacillales; Staphylococcus.
REFERENCE  1  (bases 1 to 48)
  AUTHORS  Dyke,K.G.H., Curnock,S.P., Golding,M. and Noble,W.C.
  TITLE    Cloning of the gene conferring resistance to mupirocin in
           Staphylococcus aureus
  JOURNAL  FEMS Microbiol. Lett. 77, 195-198 (1991)
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="Staphylococcus aureus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:1280"
                     /clone="pOX301"
                     /plasmid="pJ2947"
     gene            1. .48
                     /gene="mupirocin resistence gene"
     CDS             <1. .>48
                     /gene="mupirocin resistence gene"
                     /note="E.coli isoleucyl tRNA synthetase homologue"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA42079.1"
                     /db_xref="GI:46622"
                     /db_xref="GOA:P41368"
                     /db_xref="SWISS-PROT:P41368"
                     /translation="RVEEVIDVWFDSGSMR"
ORIGIN

  Query Match             62.0%;  Score 12.4;  DB 1;  Length 48;
  Best Local Similarity   92.9%;  Pred. No. 3.2e+05;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 ATCTATGTTTGGTT 20
              ||| ||||||||||
Db         16 ATCGATGTTTGGTT 29


Search completed: September 23, 2004, 16:21:00
Job time : 1274 secs
```